AD_____

Award Number DAMD17-93-J-3003

TITLE: "Wavelet Representations for Digital Mammography"

PRINCIPAL INVESTIGATOR: Andrew F. Laine, D.Sc.
Fred Taylor, Ph.D.

CONTRACTING ORGANIZATION: University of Florida
Gainesville, Florida 32611

REPORT DATE: April 1999

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | April 1999 | Final (16 Nov 92 - 15 Mar 99) |

4. TITLE AND SUBTITLE
Wavelet Representations for Digital Mammography

5. FUNDING NUMBERS
DAMD17-93-J-3003

6. AUTHOR(S)
Andrew F. Laine, D.Sc.
Fred Taylor, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Florida
Gainesville, Florida 32611

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We report on a receiver operating characteristics (ROC) study focusing on dyadic wavelets for enhancement of mammographic features in digitized mammograms. The enhancement protocol was based on multiscale expansions and non-linear enhancement functions described previously in our annual reports. In this study, dyadic spline wavelet functions were used together with a sigmoidal non-linear enhancement function. We designed a prototype test bed interface and performed a ROC study with three radiologists specialized in mammography. Data was obtained from the national mammography database of digitized radiographs from the University of South Florida.

An initial analysis of the data counted the number of false-positives and true-positives in each of two cases: Enhanced and Original. Lesions with a LOC greater or equal 3 were considered malignant. The average TPF was found to be 0.667 with enhancement, and TPF = 0.569 without enhancement. This observed increase in sensitivity is encouraging, but accompanied by a slight increase in the fraction of false-positives (0.222 compared to 0.178). However, when the analysis of the data only focused on micro-calcifications alone, we observed a TPF = 0.417 with enhancement compared to TPF = 0.222 without enhancement. No increase or decrease in FPF was observed. This finding reinforces our hypothesis that feature specific enhancement protocols are indeed useful for visualizing subtle mammographic features.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
57

16. PRICE CODE

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

# Foreword

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

__✓__ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals, "prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Andrew Lawrence 10/12/99

PI - Signature Date

# Table of Contents

# Introduction

This document is the final report of research project entitled "Wavelet Representations for Digital Mammography," sponsored by the Breast Cancer Research Program of the Department of Defense U.S. Army Medical Research and Material Command. It describes experimental methods, assumptions, procedures and results of *Phases IV* and *V* of the Statement of Work, as revised July 1997. Accomplishments relative to completion of *Phase IV* , *"Visualization Requirements for Evaluation Studies"* and *Phase V "Perform a Retrospective Study on Existing Local and National Mammography Databases,"* are summarized below.

## EXECUTIVE SUMMARY

In the final Phases of this project, we carried out a receiver operating characteristics (ROC) study focusing on dyadic wavelets for enhancement of mammographic features in digitized mammograms. The enhancement protocol was based on multiscale expansions and non-linear enhancement functions described previously in our annual reports. Specifically, in this case dyadic spline wavelet functions were used together with a sigmoidal non-linear enhancement function. In this final phase, we designed a prototype test bed interface and performed a ROC study with three radiologists specialized in mammography. Data was obtained from the national mammography database of digitized radiographs from the University of South Florida.
Susan Smith, M.D. along with three additional radiologists specializing in mammography, of the Breast Imaging Center at Presbyterian Hospital participated in the preliminary ROC study described below. All three mammographers participating in this study had a previous background in CAD systems evaluations, metrics for image quality [9] and ROC studies.

### 1. Selection of Cases

To measure the benefits of diagnosing digitized mammograms with enhancement through multiscale expansions, this study focused on dense mammograms, i.e. mammograms of density 3 and 4, which are the most difficult cases in screening. In general, the enhancement protocol aimed at improving the detection and localization of mammographic features, such as microcalcifications, masses, and spicular lesions without introducing "false-positives".
To compare the performance of radiologists with and without using the enhancement tool, two groups of 30 cases each were presented. Each group contained 15 cases of cancerous and 15 cases of normal mammograms. As mentioned above, a national mammography database of the University of South Florida provided "ground truth" (mostly through biopsy) for the selected cases. The selection was carried out carefully under the guidance of Dr. Smith, in order to find challenging cases of the same difficulty for each group. Images showing metal markers ("bibis") to indicate suspicious regions of breast tissue were avoided as well as obvious malignancies.

### 2. Display Setup and Software

Images from the mammography database were digitized from film at the resolutions of 40 to 50 μm. Image widths vary between 2000 and 3000 pixels, and image heights from 4000 to 5900 pixels. Depending on the scanner utilized for digitization the contrast resolution was either 12 bits or 16 bits per pixel resulting in large amounts of data. The files were stored in RAW binary format.
The graphical user interface (GUI) developed for this study was written in Visual C++ 6.0, whereas the code for the wavelet expansion and image reconstruction was written in native "C" to speed performance. To handle the large amounts of data and to provide the diagnosing radiologist with as much information as possible all four views (right and left medial-lateral (RMLO, LMLO) and right and left cranial caudal (RCC, LCC)) of a case were loaded into memory and displayed as downsampled images. Downsampling was still necessary to fit the images on the screens. Two high-resolution MegaScan monitors with a screen size of 2048 by 2560 were used. The four views were aligned to help the radiologist to look for asymmetries. In addition, one view could be

selected. A viewport displayed a ROI at full resolution from a mammogram in this view. The size of the viewport could be 512 by 512, 1014 by 1024 or 2048 by 2048. The center of the ROI was determined through a mouse pointer in the chosen view. Thus, the original mammogram could also be viewed through the viewport, if desired. More importantly, suspicious areas could be captured in the viewport and processed through enhancement via multiscale expansion. The number of subbands of the expansion could be adjusted by the user as well. After selecting a ROI, processing was applied to the corresponding matrix. The image was decomposed onto dyadic wavelet basis functions yielding wavelet coefficients. Coefficients were modified by a sigmoidal non-linear enhancement function, and the image was reconstructed from modified coefficients in nearly real-time.

For each subband of the multiscale expansion each of the two parameters could be adjusted trough sliders. On release of the slider button reconstruction was triggered, and a resulting image presented in a new window. Reconstruction of a 512 by 512 matrix for five levels of decomposition (5 subbands) took 5 to 6 seconds. A four subband reconstruction took on average 4 to 5 seconds. However, this could be reduced to achieve true real-time performance, by optimizing the program. Results of enhanced images could be saved together with its corresponding downsampled view, where the position of each ROI was recorded.

The enhancement protocol was run on an IBM IntelliStation Z Pro Professional Workstation Type 6865. This machine has two Intel Pentium II Xeon microprocessors (450 MHz), 512MByte of RAM and is equipped with 36 GByte of hard disk space. Windows NT 4.0 was the operating system.

## 3. Paradigm of the Preliminary Study for Evaluation of Enhanced Mammograms.

The procedure followed by each radiologist is described below:

- Without Enhancement:

The radiologist made a diagnosis based only on the four original displays and the viewport. No processing of ROIs was allowed.

- With Enhancement:

The radiologist selected a Region of Interest (ROI) on one of the views. Four levels of scales were computed. No enhancement function was applied initially. The result of the multiscale enhancement on the ROI was displayed in a new window. The radiologist then evaluated the quality of the enhanced ROI and adjusted the equalizer sliders of a channel to improve the visual quality of the suspicious region. Once he/she was satisfied with the visual result or if he/she judged that total satisfaction could not be achieved with the given tool, he/she made a diagnostic decision.

A diagnosis included specifying all lesions found and assigning a BI-RAD scale to each breast and the case.
In addition, the radiologist was asked to choose a level of confidence (LOC) in a positive diagnosis, i.e. cancer is present, on an integer scale from 1 (total confidence that there are no malignant lesions) to 5 (total confidence that there is a malignant lesion). The value for the level of confidence was used in the analysis of data to decide whether a lesion was classified as malignant or not.

## 4. Results of the Preliminary Study

An initial analysis of the data counted the number of false-positives and true-positives in each group of cases. To consider a lesion as being diagnosed as malignant or benign, the LOC value was thresholded [32]. This threshold influences the shape of the ROC curve and its interpretation. In general, any enhancement protocol should increase sensitivity, i.e. fraction of true-positives (TPF), without decreasing specificity, i.e. essentially without increasing the fraction of false-positives (FPF).

If the threshold for the level of confidence was chosen to be 3, meaning that lesions with a LOC greater or equal 3 were considered as malignant, then the average TPF was found to be 0.667 with enhancement, and TPF =

0.569 without enhancement. This increase in sensitivity is encouraging, but was accompanied by a slight increase in the fraction of false-positives (0.222 compared to 0.178). The latter is not surprising, *since the applied enhancement protocol only used dyadic spline wavelets* with a non-linear sigmoidal enhancement function, which is not the optimal choice for all types of lesions. As suggested in the original proposal of this project, dyadic wavelet expansions are best used to enhance microcalcifications. If the analysis of the data only focuses on microcalcifications, then we observed TPF = 0.417 with enhancement compared to TPF = 0.222 without enhancement. No increase or decrease in FPF was noticed. This observation reinforces our hypothesis that feature specific enhancement protocols are indeed useful for visualizing subtle mammographic features.

## 5. Relavance to Statement of Work (Revised July, 1997).

These efforts correspondence to the goals and tasks identified in Phase IV – *Visualizualization requirements for evaluation studies*, and Phase V – *Peform a retrospective study on existing local and national mammography databases*.

# Body

## A. Enhancement Protocol

### Contrast Enhancement via Multiscale Expansions: A Short Overview

We summarize below, our previous use of overcomplete multiscale representations for adaptive contrast enhancement of mammograms. Critically sampled multiscale representations have been successfully used for compression purposes and signal analysis, but are not suitable for detection and enhancement tasks because of aliasing effects introduced during downsampling of the analysis [1], [2]. Overcomplete representations avoid such aliasing artifacts and offer the desirable property for image enhancement, of being shift invariant [3], [4]. Indeed, this property will ensure that the spatial location of any mammographic finding within an image will be preserved across all levels of scale. Note that the transform coefficient matrix size at each scale remains the same as the spatial resolution of the original image, since there is no downsampling across each level of analysis.

Overcomplete multiscale analysis and reconstruction algorithms using dyadic scales previously developed in [5], [6], and [7] and were used as an initial choice of analysis function for our preliminary study of the enhancement protocol. The implementation has been carried out using several lowpass filters and highpass filters with defined frequency support. Each level corresponds to a set of filters and two branches: one for the filtered image and one for the image at the previous level minus the filtered image of the current level. This cascade of filters enables successive decompositions of an original image into finer and finer levels of analysis, and estimation of the image into coarser levels in reconstruction. Figure 1 below, illustrates this filter bank structure. In practice, a gain function modifies the matrices of coefficients that have been isolated by the filters at each level and may boost coefficients at some scales and/or attenuate others. The framework for the high-speed execution of enhancement processing by an analysis-reconstruction algorithm is illustrated in Figure 1.

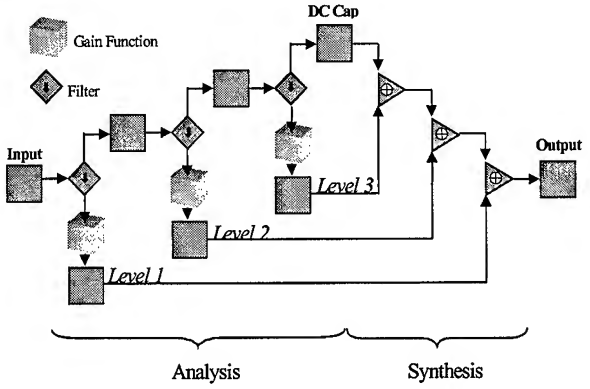



**Figure 1: Multiscale analysis with non-linear gain function. (a) Filter bank implementation, (b) Example of the processing of a ROI of a Chest radiograph. Normalized pixel intensity along a scan line that crosses a nodule is displayed for both the original and the processed image.**

The modified matrices of coefficients are simply "plugged in" during reconstruction producing a "focused" subband enhancement. As shown above, the gain function can be implemented independently of a particular set of filters and easily incorporated into a filter bank to provide the benefits of multiscale enhancement [8], [9].

### Fast Implementation

Similar to orthogonal and biorthogonal discrete wavelet transforms [10], the discrete dyadic wavelet transform can be implemented within a hierarchical filtering scheme. Let an input signal $x(n)$ be real, $x(n) \in l^1(Z), n \in [0, N-1]$ (i.e., x(n) is supported on the index interval [0, N-1]) and let $X(\omega)$ be its Fourier transform. Depending on the length of each filter impulse response, filtering an input signal may be computed

either by multiplying $X(\omega)$ by the frequency response of a filter or by circularly convolving $x(n)$ with the impulse response of a filter. Of course, such a periodically extended signal may change abruptly at the boundaries causing artifacts. A common remedy for such a problem is realized by constructing a mirror extended signal

$$x_{me}(n) = \begin{cases} x(-n-1) & \text{if } n \in [-N,-1] \\ x(n) & \text{if } n \in [0,N-1] \end{cases}$$

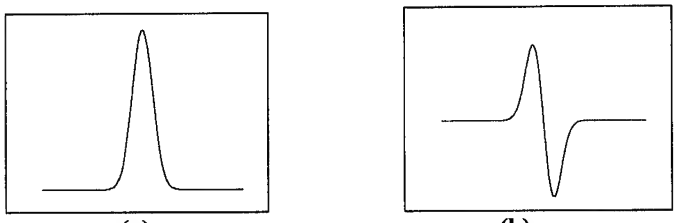
where we chose the signal $x_{me}(n)$ to be supported in [-N,N-1]. In [8] it is shown how a mirror extension is particularly elegant solution in conjunction with symmetric/antisymmetric filters.
The optimized circular convolution described in [8] has been implemented in native "C" to speed up performance for multiscale decomposition and image reconstruction. This algorithm was incorporated into the graphical user interface (GUI) developed during this phase of the study.

The benefits of one specific enhancement protocol were investigated during the academic year September 1998 to May 1999. As described in the statement of work (July 1997), we envision developing feature specific enhancement protocols for each type of lesion. Each protocol would include a multiscale expansion of a mammogram with a specific basis and an associated non-linear enhancement function that best revealed information in a mammogram for this type of lesion, e.g. microcalcifications. For the study described in this report, a dyadic Spline wavelet function was used as the basis, and a non-linear sigmoidal function was applied as the enhancement function. Both are described next in greater detail below.

**Dyadic Spline Wavelet Algorithm**

The wavelet transform of a signal $f(x)$ at a scale $s$ and position $x$ is defined by $W_s f(x) = f * \psi_s(x)$, where $\psi_s(x) = \frac{1}{s}\psi(\frac{x}{s})$ and $\psi(x)$ is the wavelet function whose average is zero.

To allow fast numerical implementation of discrete wavelet transforms, Mallat and Zhong [11] introduced a dyadic wavelet where the scale parameter varies only along the dyadic sequence $\{2^j\}$, with $j \in Z$. The 2-D dyadic wavelet transform partitions plane orientations into two bands. This means that there are two channels of analysis along the orthogonal $x$ and $y$ direction. The wavelet transform of the 2D signal $f(x,y)$ at the scale $2^j$ has two components defined by: $W^1_{2^j} f(x,y) = f * \psi^1_{2^j}(x,y)$ and $W^2_{2^j} f(x,y) = f * \psi^2_{2^j}(x,y)$ , with $\psi^d_{2^j}(x,y) = \frac{1}{2^{2^j}}\psi^d(\frac{x}{2^j},\frac{y}{2^j})$, *(d=1,2)*. In this final phase of the project, we used the particular quadratic spline wavelet function defined by Mallat and Zhong in [11] of compact support and continuously differentiable. It is the derivative of a smoothing cubic spline function as displayed in Figure 2 below.



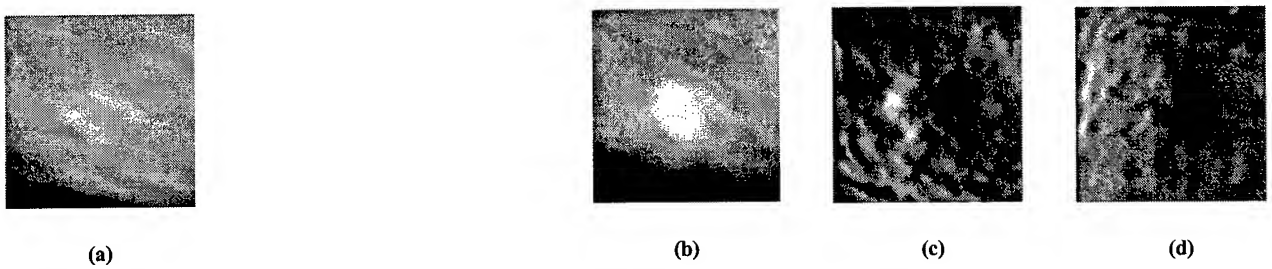
(a) (b)

**Figure 2: (a) Spline smoothing function, (b) Quadratic spline wavelet of compact support defined as the derivative of the smoothing function.**

In this context, the wavelet transform $W^d_{2^j} f$ of the signal $f$ is proportional to the derivative of the signal smoothed at the scale $2^j$. The coefficients of modulus maxima detection is then equivalent to an adaptive sampling that finds a signal variation points in the two orthogonal directions $x$ and $y$.

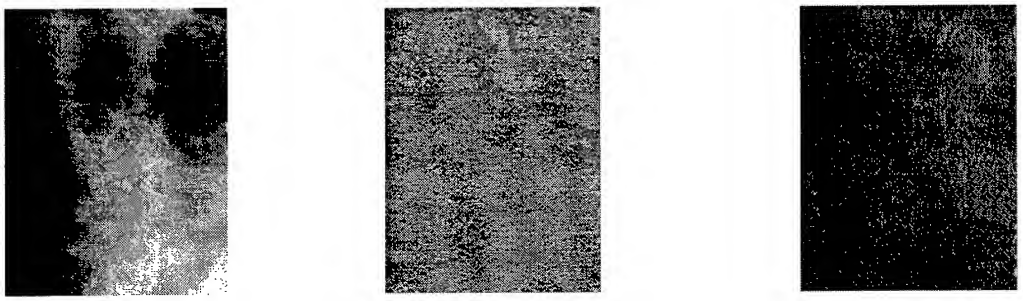
As images represent finite energy signals measured at a finite resolution, we cannot compute the wavelet transform at scales below the limit set by this resolution. We applied this analysis at integer scales varying from 1 (original signal) to the limit imposed by the acquisition resolution (digitizer sampling rate).

Figure 3 shows an example for one level of an overcomplete wavelet decomposition of a spiculated mass, and Figure 4 exhibits selecting of microcalcifications as wavelet coefficients at the finest dyadic scale.



(a) (b) (c) (d)

**Figure 3: Level 5 of an overcomplete dyadic wavelet decomposition of a spiculated mass. (a) Original image. (b) Approximation image. (c) Horizontal details. (d) Vertical details.**



(a) (b) (c)

**Figure 4: (a) Original ROI with microcalcifications. Horizontal (b) and vertical (c) dyadic wavelet coefficients.**

## Brushlet multiscale functions

During the past year, in addition to dyadic Spline wavelets we investigated [12] the brushlet basis introduced by F. Meyer and R. Coifman in [13] in 1997 for efficient compression of texture. The brushlet functions are complex valued, well localized in the frequency domain. Their construction is based on a windowed Fourier transform of the Fourier transform of the image. The projection on the orthonormal basis of brushlet functions provides a decomposition of the image along distinct orientations. We are optimistic that we can to take advantage of the special characteristics of the brushlet functions in the context of the continuation of the work reported here.

The general scheme of the analysis performed by the brushlet is the following. Let us call $f$ a given signal and $\hat{f}$ its Fourier transform. We can project $\hat{f}$ on the brushlet basis, $\hat{f} = \sum_n \sum_j \hat{f}_{n,j} u_{n,j}$ with $u_{n,j}$ the brushlet basis function and $\hat{f}_{n,j}$ the brushlet coefficients as described in [13]. The Fourier transform domain of the signal is

divided into subintervals $[a_n, a_{n+1}]$ of size $l_n$. For each interval indexed $n$, the signal $\hat{f}$ is projected on $u_{n,j}$, with $j = 0, \frac{1}{l_n}, \ldots, \frac{l_n - 1}{l_n}$.

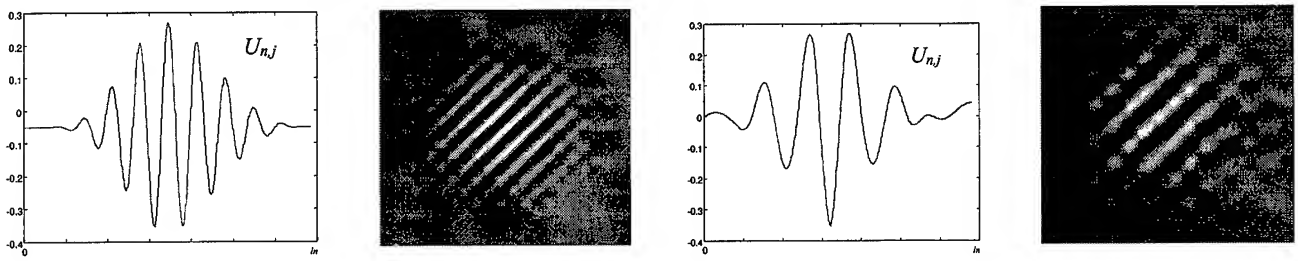
The brushlet function is defined as $u_{j,n}(x) = b_n(x - c_n)\frac{e^{-2i\pi j(x-a_n)/l_n}}{\sqrt{l_n}} + v(x-a_n)\frac{e^{-2i\pi j(x+a_n)/l_n}}{\sqrt{l_n}} + v(x-a_{n+1})\frac{e^{-2i\pi j(2a_{n+1}-x-a_n)/l_n}}{\sqrt{l_n}}$ on the interval $[a_n - \varepsilon, a_{n+1} + \varepsilon]$, with $\varepsilon$ the overlap parameter between two adjacent intervals. The widow function $b_n$ and the "bump" function $v$ define the length of the support of $u_{n,j}$ as illustrated in Figure 5 below.


(a) (b)

**Figure 5: (a), Windowing function bn, and bump function v defined on the interval [an-ε, an+1+ε]. (b), Real part of brushlet basis function uj,n.**

By applying the inverse Fourier transform, we have a decomposition of $f$, $f = \sum_n \sum_j \hat{f}_{n,j} w_{n,j}$ on the orthonormal basis $w_{n,j}$, inverse Fourier transform of $u_{n,j}$. The $w_{n,j}$ functions are defined as:

$w_{n,j}(x) = \frac{1}{\sqrt{l_n}} e^{2i\pi a_n} e^{i\pi l_n x} \left( (-1)^j \hat{b}_n\left( x - \frac{j}{l_n} \right) - 2i \sin(\pi l_n x) \hat{v}\left( x + \frac{j}{l_n} \right) \right)$. The parameter $l_n$ appears as a scaling factor of the analysis and $j$ is the translation index of the brushlet, so that $w_{n,j}$ has an expression similar to a wavelet. The phase of the function encodes the orientation of the brushlet pattern in the 2-D case as illustrated in Figure 6.

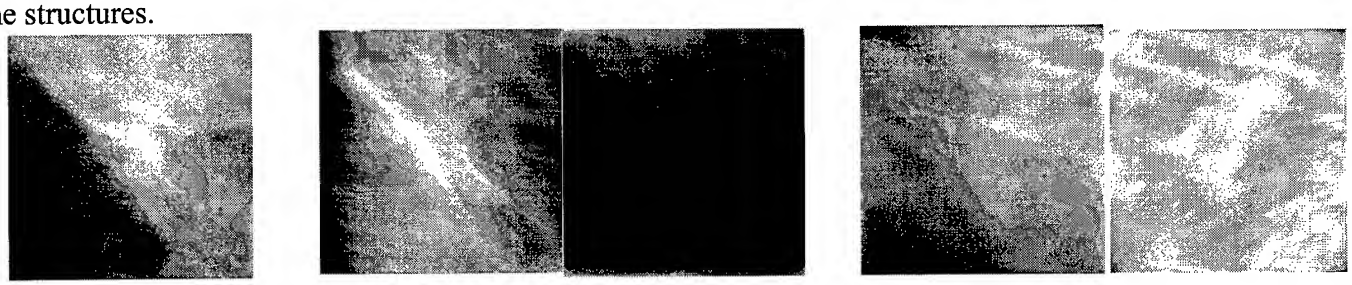



a.1 a.2 b.1 b.2

**Figure 6: (a.1-a.2) real part of 1D brushlet basis function, (b.1-b.2) real part of 2D brushlet basis function for two different scale parameter value $l_n$s and the length of window function $b$ in 1-D and size of the quadrants in the Fourier plane in 2-D.**

The projection of $\hat{f}$ on $u_{n,j}$ is efficiently implemented by the folding technique and Fourier transform. With a division of the image into four quadrants, the decomposition on $u_{n,j}$ provides four sets of coefficients showing the texture with patterns oriented along the directions $\frac{-3\pi}{4}, \frac{-\pi}{4}, \frac{\pi}{4}$, and $\frac{3\pi}{4}$. An arbitrary number of orientations are possible to construct.

Meyer has shown that these bases can lead to efficient compression of richly textured images. We believe that such basis can be applied to mammograms for directional feature enhancement, texture analysis and segmentation. Below in Figure 7 we illustrate the ability of the brushlet to decompose textures into distinct directions within a selected region of interest of a mammogram containing a spiculated mass oriented in -45° direction. The modulus of the coefficient for analysis in +45° and -45° shows strong values in the orientation direction of the mass and flat low values in the orthogonal direction. Selective amplification of the coefficients in the -45° direction and attenuation of the coefficients in +45° will enhance the spicular lesion and details of its fine structures.



(a) (b.1) (b.2) (c.1) (c.2)

**Figure 7: (a). Original ROI in mammogram with spicular lesion. (b.1-b.2) Brushlet coefficients in $\frac{-\pi}{4}$. (c.1-c.2) Brushlet coefficients in $\frac{\pi}{4}$.**

We believe that the ability of the brushlet functions to decompose the signal under different texture orientations is particularly well suited for the enhancement of spicular subtle lesions in the mammograms. Adjustment of the scale parameter $l_n$ modifies the resolution of the analysis in terms of texture orientation and oscillation frequency.

Indeed, the idea of building a specialized detector for spicular lesions with brushlet functions is very promising. We hope to continue this direction through additional support from the National Institute of Health (NIH) and the US Army Breast Cancer Research program.

## Non-Linear Enhancement Function

The enhancement process modifies the analysis coefficients within distinct subbands. This is illustrated in Figure 8 below.

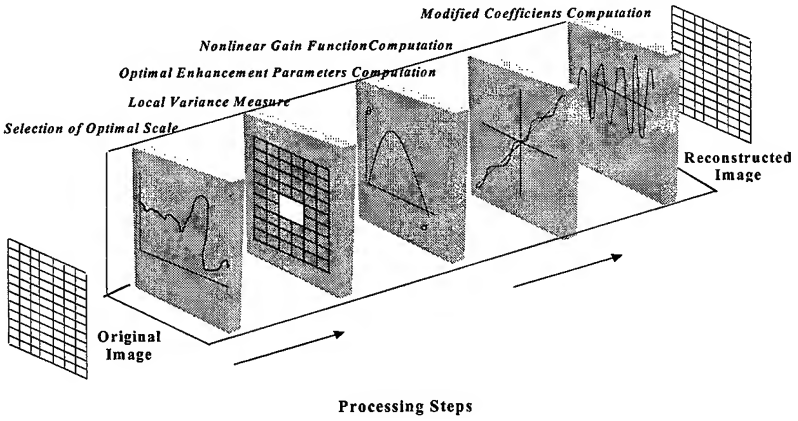



**Figure 8: Overview of multiscale enhancement protocol.**

Modification of selected analysis coefficients within a certain scale can make more obvious indiscernible or barely seen features [14]. A framework for contrast enhancement was achieved by applying a non-linear function to multiscale coefficients. This operation resulted in attenuation or local increasing of coefficients. Enhancement or gain functions must be cumulative and monotonically increasing in order to preserve the

original information in the image and to avoid artifacts [6]. Figure 9(a) provides a very simple example of a piecewise linear gain function. The parameter $w_{ij}$ represents the modulus of a multiscale coefficient. Coefficients are modified by the gain function $f(w_{ij})$. $T$ is the threshold of the function. For $\theta < 45°$ there will be an attenuation of the coefficients ($\alpha<1$), at $\theta = 45°$ we have the identity function ($\alpha=1$). For $\theta > 45°$ there is a smooth amplification of the coefficients ($\alpha>1$) below the threshold value. The values of the two parameters, $T$ and $\theta$, determine the final shape of the gain function. Figure 9(b) displays a gain function of employing hard-thresholding for denoising. Unfortunately, These two particular examples have the disadvantage of being discontinuous at the threshold value $T$. This could result in an abnormal distribution of coefficient values in the output and may create sharp peaks on both ends of the histogram of a particular output mapping. For this reason, smoother functions, like sigmoids, are preferable and were used in this project. Figure 9(c) shows an example of such a function as described in [15].

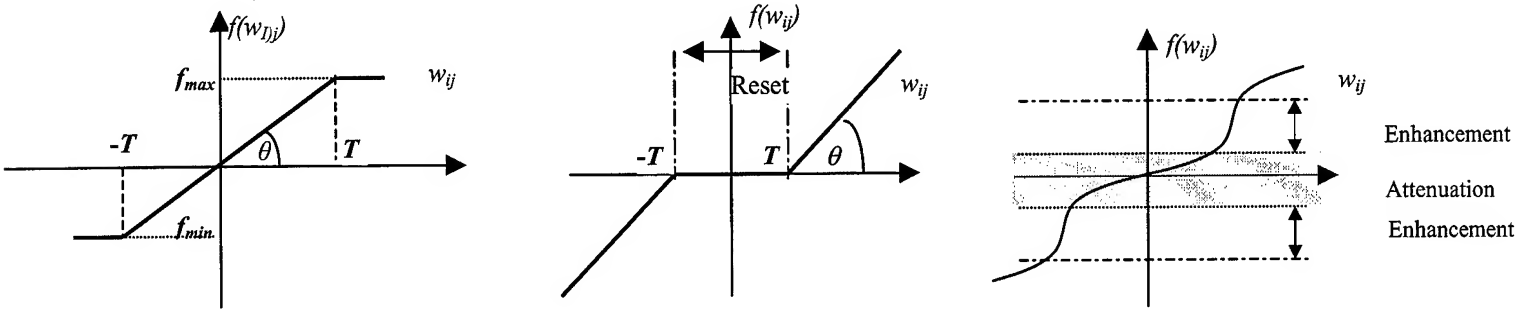



**Figure 9: (a) A simple piecewise linear enhancement function, (b) hard-thresholding, (c) .a sample non-linear enhancement function.**

The analytical formulation of the gain function as we designed it in [15], [16] is the following:

$$f(\mathrm{w}_{ij}) = a\left[sigm\left(c(\mathrm{w}_{ij} - b)\right) - sigm\left(-c(\mathrm{w}_{ij} + b)\right)\right]$$

$$a = \frac{1}{sigm\left(c(1-b)\right) - sigm\left(-c(1+b)\right)}, \quad 0 < b < 1$$

$$sigm\,(y) = \frac{1}{1+e^{-y}}$$

Parameters $b$ and $c$ control the threshold and the rate of enhancement respectively. The gain function is continuous and monotically increasing, and has a continuous first derivative. This ensures that the gain function will not introduce any new discontinuities of coefficients in the transform domain.

This particular gain function decreases the value of the coefficients in the center range of values around zero, which is equivalent to a denoising action, while it increases the values of the coefficients outside this range, equivalent to enhancement. This type of non-linear (smooth) gain function, in 'steps', offers a very rich and flexible paradigm to carry out non-linear dynamic analysis of coefficients within a specific scale [17].

There are many criteria for the selection of the enhancement function applied to the coefficients of a particular level of analysis for contrast enhancement. A preliminary goal of the phase of this project was to develop a research tool for testing enhancement functions targeted for specific mammographic features. As this process requires specialized expertise and a substantial time investment, no systematic study of the problem of associating enhancement functions with target features in mammograms has been reported in the literature.

The two parameters required for the enhancement processing are threshold and gain/attenuation. The gain function is sigmoidal and will enhance coefficients above the threshold value and decrease the coefficients below the threshold of the order of the gain amplitude.

In [16] we used quantitative information retrieved from the image to compute the threshold and the gain amplitude. Non linear estimators are signal dependent and behave differently for different realizations of each signal. In this frame of work, Johnstone and Donoho have shown that by considering the signal as deterministic, thresholding of wavelet coefficients gives a nearly optimal estimation of piecewise smooth functions [18], [19]. Selection of the threshold value was based on comparison with local variance in the transform domain. For a noisy signal of size $N$, thresholding of the wavelet coefficients with $T = \sigma\sqrt{2\ln(N)}$ where $\sigma$ is the coefficients standard deviation provides an asymptotically optimal estimator of the original signal in the mini-max sense. Soft thresholding of the wavelet coefficients performs an adaptive smoothing of the image by averaging the noisy areas and preserving or enhancing coefficients in areas of sharp transitions. Noise standard deviation can be estimated by determining the median wavelet coefficient value at the finest scale or with local discrete statistical estimation in the transform domain. Using extremely local variances leads to a very aggressive posturing of the gain function, and represents a high amount of intervention in adjusting the output, while global variance measurements were less noticeable. Superiority of either method depends on the screening protocol used by the radiologist and the kind of analysis to be performed. For example, fine microcalcifications represent high frequency information of the image. We would expect the local variance for such a feature will be high with a selected ROI. Consequently, smooth amplification of coefficients within this particular spatial frequency (in combination with possibly decreasing the information of other spatial frequencies) will enhance these features of interest. Similar analysis can be done to enhance low spatial frequency features such as masses. Since the computation of the threshold and the gain function use data dependent information such as noise, standard deviation and local coefficient variance, digital and digitized radiographs acquired under different imaging conditions are processed differently. Intrinsic properties of the radiograph are incorporated in the setting of the parameters so that enhancement is adaptively optimized to each mammogram processed.

## B. Development of a Graphical User Interface (GUI)

### Motivation

Running an enhancement algorithm in a batch mode might be sufficient for research purposes. However, adjustment of parameters ties to a data dependent enhancement function is slow because of the repeated need to decompose and reconstruct from modified coefficients. A more desirable situation is to observe the results of modified multiscale coefficients and to continue the enhancement procedure, until results are visually satisfactory or the decision is made that no further improvement can be achieved. In addition, with introducing fixed enhancement protocols into a clinical screening paradigm, the algorithm must be simple, fast, and user-friendly, i.e. usage of the algorithm should be familiar to the radiologist and intuitive. Since each radiologist may have preferences with respect to contrast in mammograms, it must be possible to adjust parameter settings to those preferences. Thus, a graphical user interface was designed to facilitate carrying out a such a studiy and to create a software prototype, whose successors might find entrance into clinical screening. We call this application a "test bed" softcopy display tool.
The test bed softcopy display tool provided our research team a means to carry out rapidly, experimental studies for sigmoidal enhancement function and compute optimal values for threshold and gain values using information extracted from selected ROIs. It enabled quick comparison of results and made feasible a methodical examination with regard to measuring image quality. The first version of this research software was employed for a ROC study, which included four radiologists from the Columbia-Presbyterian Medical Center as reported in Section C of this report.

Another reason for testing user-interactive enhancement techniques in a clinical environment stemmed from the fact that New York Presbyterian Hospital, as well as others throughout the country, is undertaking an enterprise wide reorganization. The Department of Radiology is eliminating film support from daily practice and screening diagnosis for MRI and CT. In house diagnosis will be performed on soft copy display starting in July 1999. The Breast Imaging Center will not suppress film support, because of existing limitations of softcopy display. Nevertheless, in a screening environment, integration of advanced software tools to improve image quality and the specificity of findings without discarding information is of critical importance. There remains a crucial need for the development of soft copy display tools that allow the radiologist to preserve or improve his/her diagnostic performance in the context of a daily routine screening in a clinical environment. Radiologists will be confronted with new visualization technologies and new working tools redefining screening protocols. Because of the current limitations of hardware in display resolution, we believe that enhancement of mammograms will allow mammographers use these new techniques and possibly improve at the same time their confidence and diagnostic performance. The opportunity to develop a CAD tool in this context is unique and bears a potential to orient the directions of research in this field and move digital mammography forward.

### Design and Implementation

The graphical user interface (GUI) developed for this study was written in Visual C++ 6.0. This particular development environment was chosen to take advantage of already predefined classes wrapping parts of a GUI, such as sliders and dialog boxes. Moreover, the code for the wavelet expansion and image reconstruction that was written in native "C" to speed up performance could be incorporated and executed in this environment without major modifications, thus shortening development time. Some of the guidelines and considerations for the design and implementation of the GUI are described next.
The prototype test bed interface was primarily designed to process raw 16-bit data (image files without header). Data was obtained from the national mammography database of digitized radiographs from the University of South Florida. We have the complete database of digitized mammograms (stored on twenty-one 8mm tapes). Our database contained 586 selected cases of malignant lesions, biopsy proven, and 437 cases of normal breasts. More specifically, different types of lesions are represented in the following proportions: 100 round and oval malignant masses, 216 spicular lesions and 248 microcalcifications. The quality of the mammograms

varied. 559 cases of dense breasts (density of 3 and 4) with 266 normals and 293 cancerous, referred by radiologists as the most challenging cases, are included in the database.

Images from the mammography database were digitized from film at the resolutions of 40 to 50 μm. Image line length vary between 2000 and 3000 pixels, and number of rows from 4000 to 5900 pixels. Depending on the scanner utilized for digitization the contrast resolution was either 12 bits or 16 bits per pixel resulting in 15-50 megabytes per file.

To handle the large amounts of data and to provide the diagnosing radiologist as much information as possible all four views (right and left medial-lateral (RMLO, LMLO) and right and left cranial caudal (RCC, LCC)) of a case were loaded into memory and displayed as downsampled images. Downsampling was necessary to fit the images on screen, consisting of two high-resolution MegaScan monitors each with a screen size of 2048 by 2560. The four views were aligned to assist the radiologist to look for asymmetries. In addition, one view could be selected, and a viewport displayed a cursor selected ROI at full resolution from a selected mammogram. The size of the viewport could be chosen as 512 by 512, 1014 by 1024 or even 2048 by 2048. The center of the ROI was determined through the mouse pointer in a chosen window. Thus, the original mammogram could also be viewed through the viewport, if desired. More importantly, suspicious areas could be captured in the viewport and processed through enhancement via multiscale expansion. The user could adjust the number of subbands of the expansion as well. After selecting a ROI processing was applied. The image was decomposed onto dyadic wavelet basis functions yielding wavelet coefficients. Coefficients were modified by a sigmoidal non-linear enhancement function, and the image was reconstructed from these modified coefficients in nearly real-time.

As mentioned in Section A of this report the shape of the enhancement function can be changed through modification of the two parameters gain and threshold. For each subband of the multiscale expansion each parameter could be adjusted trough sliders (see Figure 10(b)). On release of the slider button reconstruction was "triggered", and a resulting image presented in an output window. Reconstruction of a 512 by 512 matrix for five levels of decomposition (5 subbands) took 5 to 6 seconds, for four subbands, reconstruction time shortened to 4 to 5 seconds. During our ROC study the application was executed in a double-buffering mode. The application was executed twice to reduce waiting time for the loading of images. Since the total amount of data to be loaded into memory for one case amounted to up to 200 MByte, it took up to 40 seconds to finish. To avoid idle times for the diagnosing radiologist, one case was loaded in the background, while she/he worked on one previously uploaded. All code was compiled to maximize speed. Reconstruction times $t_{recon}$ for different sizes of the ROI and different number of levels of analysis are given in Table 1. However, reconstruction time can be further reduced to achieve true real-time performance, by employing faster algorithms.

After processing, results of enhanced images could be saved together with its corresponding downsampled view, where the position of the ROI was marked. This was necessary to be able to evaluate a particular diagnosis for each case in comparison with the "ground truth" provided in the database. For the same case, different views and multiple ROIs out of the same view could be selected for processing. Hence, all suspicious areas in a case could be carefully examined.

| Size of Region of Interest (ROI) | $t_{recon}$ for 4 Levels of Analysis | $t_{recon}$ for 5 Levels of Analysis |
|---|---|---|
| 512 x 512 | 4-5 seconds | 6-7 seconds |
| 1024 x 1024 | 19-20 seconds | 24-25 seconds |

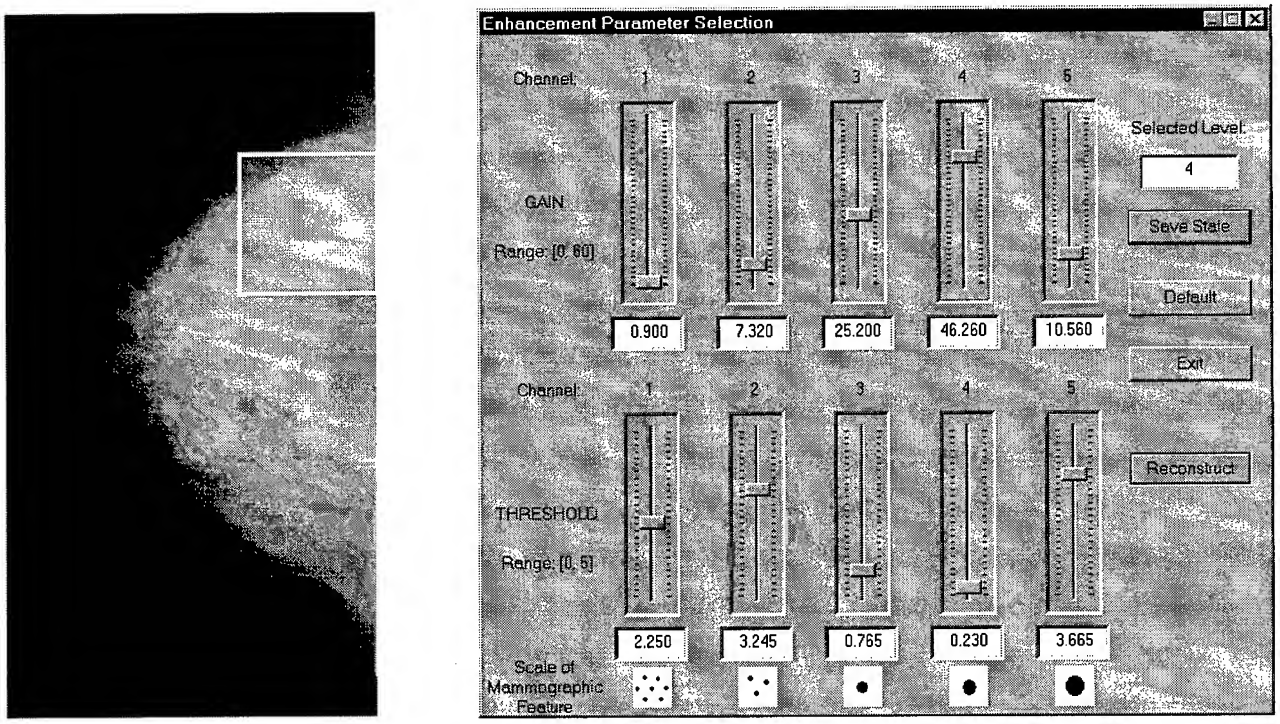
**Table 1: Reconstruction times $t_{recon}$ for two levels of analysis and two sizes of ROI.**

The enhancement protocol was run on an IBM IntelliStation Z Pro Professional Workstation Type 6865. This machine has two Intel Pentium II Xeon microprocessors (450 MHz), 512 MByte of RAM and is equipped with 36 GByte of hard disk space. Windows NT 4.0 service pack 4 was the operating system.

Figure 10(b) shows the test bed interface as an illustration of the type of tool constructed in the preliminary study of this project. For internal research and development, optimal enhancement parameters will later be

computed with information extracted from selected ROI's. Interactive (real-time) enhancement was accomplished via sliders shown in the graphical user interface (GUI). The enhancement operation relied on the optimality of parameters derived from their mathematical models and on the strategy employed for the type of enhancement applied to each subband of coefficients (amplification, preservation or diminution). Selected subband coefficients at a particular level could be strongly suppressed by choosing large thresholds (> 2) and small gains (< 1), which can be desirable for the elimination of (structured and acquisition) noise, or normal benign anatomical structures. A later version of this tool will allow display the histogram of analysis coefficients at a particular level and visualize the coefficients at any level.

We believe that these options provided sufficient flexibility for identifying feature specific enhancement protocols. Since the size of digital mammograms can be quite large, a ROI (fixed at either 512 x 512 or 1024 x 1024) within the original image is chosen to avoid the computation over region that do not contain suspicious areas. This is also shown in Figure 10, where part (a) exhibits an original digitized mammogram with a 512 x 512 ROI that contains a possible mass. Figure 10(c) and Figure 10(d) display this ROI before and after enhancement via non-linear modification of multiscale coefficients.




(a) (b)

(c) (d)

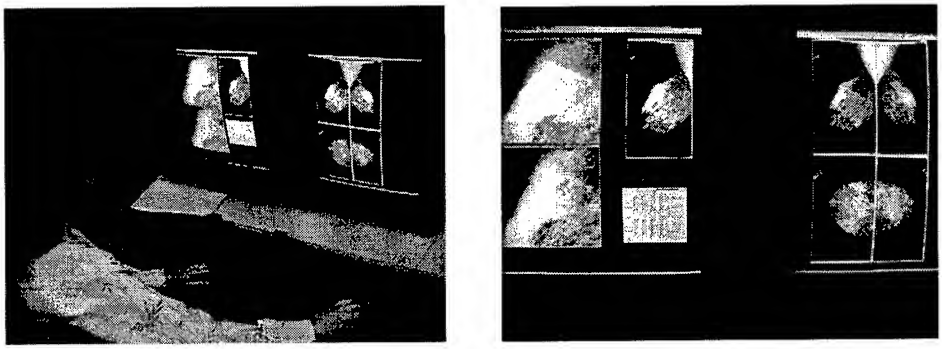
**Figure 10: (a) Original mammogram with selected ROI containing a mass, (b) Test bed interface menu, (c) Original ROI, and (d) Enhanced ROI via subband equalization.**

**Display and Hardware Settings**

High resolution displays are needed to present mammograms in an authentic way and to explore the richness of information quantized at 16-bit per pixel (bpp) grayscale data (65536 shades of gray). To meet those conditions the IBM Intellistation workstation in our laboratory has been equipped with two Metheus P1540 Graphics controllers. These are ultra high resolution display subsystems for the PCI bus with a resolution of 2048×2560 pixels each, a digital-to-analog converter (DAC) capable of 1024 shades of gray, real time window leveling. With the Metheus framebuffers, an extended hardware palette of nearly 16,000 entries can be accessed through special C++ function calls that were part of a library provided to us as developers for BARCO/Metheus. These functions wrapped *DirectDraw* functionality provided by Microsoft to obtain direct access to the video framebuffer and to take advantage of advanced display capabilities. Please see attached letter from BARCO/Metheus that certifies that our research group is an official member of the BARCO/Metheus Software Developer's program, which allowed our group to have access to the source code used for display programming. Using these library functions, the extended palette was loaded with a ramp of 4096 shades of gray corresponding to 12-bit resolution. Images stored in 16-bit per pixel format, were rescaled to 12 bpp, if necessary (most of the mammograms were digitized at a resolution of 12 bpp), and then displayed at full resolution. Direct access to the video framebuffer also sped up the display process useful for updating and refreshing the different views on the screen.

Two high-resolution MegaScan monitors were attached to a single workstation providing dual headed display on a single logical frame buffer or virtual desktop of 4000×2048 pixels, respectively with Windows NT 4.0. To ensure the accurate depiction of the same image quality on both screens, a Metheus P1500 luminance photometer was used. It recognizes the 1024 shades of gray displayed by a monitor and has a range of 0-450ft-L. Both monitors were calibrated to correct for non-linearity in through gamma correction. The Metheus display driver supports a gamma lookup table (LUT) loading function that accomplishes this. The gamma LUT can conceptually be thought to be between the palette lookup table and the actual DAC, which converts the digital luminance value into an analog luminance value (voltage) to send to the monitor. The gamma LUT was created from the real monitor luminance so that each palette intensity provided the expected linear response out of the monitor. This table was calculated by looking in the actual luminance table and finding the closest match for the desired luminance. The entire procedure can be carried out with software provided by BARCO/Metheus that measures luminance intensities, calculates a gamma LUT, which is written to file. By loading these files the non-linearity is corrected.

Figure 11 shows Dr. Koenigsberg, one of three radiologists who participated in this investigation, during the first ROC study described in this report.



(a) (b)

**Figure 11: (a) Tova Koenigsberg, M.D., using the GUI during the preliminary ROC study described above. (b) Typical screen display used during the ROC study: on the right monitor four original digitized mammograms of one case are displayed, on the left monitor, in the top-right corner is the original mammogram to be enhanced, in the down-right corner is the GUI interface, in the lower-left corner is an original ROI selected by the radiologist, in the lower-left corner a sample enhanced ROI is shown.**

Lighting conditions were controlled for the ROC study to model reading room conditions. The ambient light intensity was measured with the luminance photometer to be 12.802659 candelea/$m^2$. It is worthwhile to note that the optimality of enhancement parameters is independent of the CRT display quality and the image acquisition quality. As their computation is data driven, they are adapted to signal content and its characteristics. As our radiologists give us feedback on the quality of the enhancement, we expect to converge and adjust these initial default settings.

## C. Description of the Receiver Operating Characteristics (ROC) study

We have carried out the first receiver operating characteristics (ROC) study focusing on overcomplete dyadic wavelets for enhancement of mammographic features in digitized mammograms. The enhancement protocol was based on multiscale expansions and non-linear enhancement functions explicitly described in Section A of this report. Specifically, dyadic spline wavelet functions were used together with a sigmoidal non-linear enhancement function. The ROC study included three radiologists specialized in mammography.
The medical doctors involved in this study had a strong background in CAD systems evaluations and ROC studies. The Director of the Breast Imaging Center at Columbia Presbyterian Medical Center, Dr. Smith, assisted in the selection of cases.

### 1. Selection of Cases

To measure the benefits of diagnosing digitized mammograms with enhancement through multiscale expansions, we focused on dense mammograms, i.e. mammograms of density 3 and 4, which are the most difficult cases in screening. In general, the enhancement protocol aimed at improving the detection and localization of mammographic features, such as microcalcifications, masses, and spicular lesions without introducing "false-positives".
To compare the performance of radiologists with and without using the enhancement tool, two groups of 30 cases each were presented. Each group contained 15 cases of cancerous and 15 cases of normal mammograms. As mentioned above, a national mammography database of the University of South Florida provided "ground truth" (mostly through biopsy) for the selected cases. The selection was carried out very carefully under the guidance of Dr. Smith, in order to find rather challenging cases of similar difficulty for each group. Images showing metal markers ("bibis") to indicate suspicious regions of breast tissue were avoided as well as obvious malignancies. Due to time constraints the number of cases had to be limited for this initial study. In the future, we plan to carry out extended ROC studies with a larger number of cases and with a further optimized GUI display.

### 2. Paradigm of Diagnosis of Study

The enhancement procedure followed by the radiologist was the following:

- Without Enhancement:

The radiologist made a diagnosis based only on the four original displays and the viewport. No processing of ROIs was allowed.

- With Enhancement:

The radiologist selected a Region of Interest (ROI) on one of the views and could apply multiscale enhancement. Four levels of scales were computed. The result of the multiscale enhancement on the ROI was displayed in a new window. The radiologist then evaluated the quality of the enhanced ROI and adjusted the equalizer sliders of a channel to improve the visual quality of the suspicious region. Once he/she was satisfied with the visual result or if he/she judged that total satisfaction could not be achieved with the given tool, he/she made a diagnostic decision.

A diagnosis included specifying all lesions found and assigning a BI-RAD scale to each breast and the case.
In addition, the radiologist was asked to choose a level of confidence (LOC) for each positive diagnosis, i.e. cancer is present, on an integer scale from 1 (total confidence that there are no malignant lesions) to 5 (total confidence that there is a malignant lesion). The value for the level of confidence was used in the analysis of data to decide whether a lesion was classified as malignant or not.

## 3. ROC Data

Table 2 and Table 3 on the next two pages show the data cquired during the first ROC study. Group1 comprises the set of cases, where the radiologists were allowed to take advantage of the enhancement protocol, whereas group 2 contains those cases, where no processing could be applied. Each of the tables shows the case numbers, the case designation and total number (#) of lesions for each case according to the mammography database, and for each of the three mammographers the BI_RAD rating and level of confidence (LOC) values. The BI_RAD rating could be chosen from the standard categories 0-5, with 0 meaning that additional information for a more confident diagnosis was needed. In those cases, the radiologists were asked to also select a BI_RAD rating different from 0, if they were asked to make a diagnosis without any additional information. This number is shown in parentheses for the corresponding cases.

Both groups are sorted into actually-negative cases (normals with 0 lesions) and actually-positive cases (cancers with, at least 1malignant lesion), since this was required for subsequent data analysis.

| Group1 | (with Enhancement) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Mammographer 1 | | Mammographer 2 | | Mammographer 3 | |
| Case # | Database | DB_Total # of Lesions | BI_RAD | LOC | BI_RAD | LOC | BI_RAD | LOC |
| 2 | A_0058 | 0 | 4 | 3 | 1 | 1 | 3 | 2 |
| 5 | A_0069 | 0 | 1 | 2 | 1 | 1 | 1 | 1 |
| 6 | A_0041 | 0 | 3 | 2 | 1 | 1 | 1 | 1 |
| 7 | A_0077 | 0 | 3 | 2 | 2 | 1 | 2 | 1 |
| 9 | A_0064 | 0 | 2 | 2 | 2 | 1 | 2 | 2 |
| 13 | A_0067 | 0 | 0(3) | 2 | 1 | 1 | 0(3) | 3 |
| 15 | A_0080 | 0 | 0(3) | 3 | 2 | 1 | 2 | 1 |
| 16 | A_0089 | 0 | 3 | 3 | 1 | 1 | 1 | 2 |
| 19 | A_0062 | 0 | 2 | 2 | 1 | 1 | 2 | 1 |
| 21 | A_0057 | 0 | 2 | 2 | 1 | 1 | 0(3) | 3 |
| 24 | A_0072 | 0 | 1 | 2 | 1 | 1 | 1 | 1 |
| 25 | A_0070 | 0 | 1 | 2 | 0(3) | 2 | 1 | 2 |
| 26 | A_0068 | 0 | 1 | 2 | 1 | 1 | 2 | 1 |
| 28 | A_0039 | 0 | 3 | 2 | 1 | 1 | 0(4) | 3 |
| 30 | A_0092 | 0 | 3 | 2 | 1 | 1 | 1 | 1 |
| | | | | | | | | |
| 1 | B_3044 | 1 | 4 | 4 | 4 | 4 | 4 | 3 |
| 3 | B_3073 | 1 | 3 | 2 | 3 | 2 | 4 | 3 |
| 4 | B_3006 | 1 | 5 | 5 | 5 | 5 | 5 | 5 |
| 8 | B_3032 | 1 | 0(3) | 2 | 5 | 4 | 4 | 4 |
| 10 | B_3107 | 1 | 5 | 4 | 4 | 4 | 5 | 4 |
| 11 | C_0060 | 1 | 0(3) | 3 | 0 | 3 | 0(4) | 3 |
| 12 | B_3057 | 1 | 4 | 4 | 5 | 4 | 4 | 4 |
| 14 | B_3078 | 1 | 5 | 4 | 5 | 4 | 0(4) | 3 |
| 17 | B_3033 | 1 | 0(3) | 2 | 0 | 2 | 0(3) | 3 |
| 18 | B_3031 | 1 | 0(4) | 4 | 5 | 4 | 0(3) | 3 |
| 20 | B_3076 | 1 | 0(3) | 3 | 0 | 3 | 0(5) | 4 |
| 22 | B_3058 | 1 | 5 | 5 | 5 | 5 | 4 | 4 |
| 23 | B_3079 | 1 | 2 | 2 | 1 | 1 | 1 | 1 |
| 27 | B_3047 | 1 | 3 | 2 | 0(4) | 3 | 0(4) | 3 |
| 29 | C_0008 | 1 | 0(3) | 3 | 3 | 3 | 0(4) | 3 |

**Table 2: ROC data for three mammographers for group 1, i.e. with enhancement.**

| Group2 | (without Enhancement) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Mammographer 1 | | Mammographer 2 | | Mammographer 3 | |
| Case # | Database | DB_Total # of Lesions | BI_RAD | LOC | BI_RAD | LOC | BI_RAD | LOC |
| 3 | A_0015 | 0 | 2 | 2 | 1 | 1 | 1 | 1 |
| 4 | A_0034 | 0 | 2 | 2 | 0(3) | 2 | 0(3) | 3 |
| 5 | A_0112 | 0 | 2 | 1 | 1 | 1 | 0(4) | 3 |
| 8 | A_0020 | 0 | 2 | 2 | 1 | 1 | 2 | 2 |
| 9 | A_0003 | 0 | 3 | 2 | 1 | 1 | 1 | 1 |
| 13 | A_0030 | 0 | 2 | 2 | 1 | 1 | 0(3) | 2 |
| 15 | A_0009 | 0 | 2 | 2 | 1 | 1 | 2 | 2 |
| 16 | A_0037 | 0 | 2 | 2 | 1 | 1 | 1 | 2 |
| 17 | A_0099 | 0 | 0(3) | 2 | 1 | 1 | 2 | 1 |
| 18 | A_0116 | 0 | 0(3) | 3 | 1 | 1 | 1 | 1 |
| 21 | A_0035 | 0 | 0(3) | 2 | 0(4) | 3 | 0(3) | 3 |
| 23 | A_0018 | 0 | 2 | 2 | 1 | 1 | 1 | 1 |
| 24 | A_0022 | 0 | 2 | 2 | 1 | 1 | 0(3) | 3 |
| 27 | A_0005 | 0 | 0(3) | 2 | 0(3) | 2 | 1 | 2 |
| 30 | A_0016 | 0 | 2 | 2 | 1 | 1 | 1 | 2 |
| | | | | | | | | |
| 1 | B_3003 | 1 | 1 | 2 | 1 | 1 | 5 | 5 |
| 2 | B_3389 | 1 | 2 | 2 | 1 | 1 | 1 | 1 |
| 6 | B_3009 | 1 | 0(4) | 4 | 0(3) | 2 | 0(4) | 3 |
| 7 | C_0309 | 1 | 4 | 4 | 1 | 1 | 0(4) | 3 |
| 10 | C_0142 | 1 | 0(3) | 3 | 0(3) | 2 | 1 | 2 |
| 11 | B_3016 | 1 | 0(4) | 4 | 0(3) | 2 | 4 | 4 |
| 12 | B_3382 | 1 | 2 | 2 | 1 | 1 | 3 | 2 |
| 14 | B_3134 | 1 | 5 | 4 | 4 | 4 | 5 | 5 |
| 19 | B_3005 | 3 | 0(3) | 3 | 3 | 3 | 0(4) | 4 |
| 20 | C_0127 | 1 | 0(3) | 3 | 0(4) | 3 | 0(4) | 4 |
| 22 | C_0015 | 1 | 0(4) | 4 | 0(4) | 4 | 5 | 5 |
| 25 | B_3007 | 1 | 3 | 3 | 4 | 3 | 4 | 4 |
| 26 | B_3012 | 1 | 5 | 5 | 5 | 5 | 0(4) | 3 |
| 28 | B_3380 | 1 | 0(4) | 4 | 4 | 4 | 0(4) | 4 |
| 29 | C_0358 | 1 | 5 | 5 | 5 | 4 | 0(4) | 4 |

**Table 3: ROC data for three mammographers for group 2, i.e. without enhancement.**

## 4. ROC Analysis: General Principles

The most common method to objectively evaluate the performance of a diagnostic system or the difference in performance between two diagnostic systems is ROC analysis. It compares radiologists' image-based diagnoses with actual states of disease and health. For ROC analysis performance of a diagnostic system can be meaningfully described by the indices of "sensitivity" and "specificity", where "sensitivity" can be expressed as the true-positive fraction (TPF) and "specificity" by the true-negative fraction (TNF) of a diagnosis [20]. TPF corresponds to the fraction of cases in a study that have been diagnosed as positive (diseased) and that are actually positive, and TNF corresponds to the fraction of cases that have been diagnosed as negative (healthy) and that are actually negative. In a complimentary way, the false-negative fraction (FNF) and the false-positive fraction (FPF) can be defined as FNF = 1-TPF and FPF = 1-TPF, respectively, with a similar interpretation. Due to this dependence of these indices it is only necessary to measure one pair of indices, and frequently TPF and FPF are used. In this report we also have focussed on FPF and TPF to characterize the performance of our enhancement protocol.

In general, it is desirable for a diagnostic system to increase "sensitivity" and "specificity" or, at least to increase TPF without increasing FPF.

The underlying model for ROC analysis is the use of probability density distributions of a radiologist's confidence in a positive diagnosis for a particular diagnostic task for actually positive and actually negative patients [20]. These distributions generally have different means. It is currently accepted that based on a confidence threshold, i.e. a particular level of confidence (LOC) in a positive diagnosis, a diagnosis is considered to be positive, if it exceeds this threshold, and a diagnosis is considered to be negative, if it falls below the threshold. TPF and FPF are then calculated from the probability density distributions as areas under the curves delimited by the confidence threshold (see Figure 12 below). Changing the confidence threshold yields changes in TPF and FPF that are inversely related. If the confidence threshold is varied continuously, a curve can be generated from the pair values for TPF and FPF. Conventionally, an ROC curve plots TPF (i.e., sensitivity) as a function of FPF (i.e., 1-[specificity]). Clearly, both TPF and FPF can take values between 0.0 and 1.0. Since the curve represents all of the compromises between sensitivity and specificity that can be achieved by a diagnostic system as the confidence threshold is varied, ROC curves indicating better decision performance are positioned higher in the unit square spanned by FPF and TPF. Therefore, the area under the ROC curve $A_z$ provides a useful summary index for the inherent discrimination performance of a diagnostic system. The area $A_z$ can be interpreted as the average value of sensitivity corresponding ROC curve, if the specificity of the system is selected randomly between 0.0 and 1.0. Equivalently, $A_z$ can be considered as the average value of specificity on the ROC curve, if sensitivity is selected randomly between 0.0 and 1.0 [20].

Data for an ROC analysis is obtained by providing a set of rating categories to the radiologist, from which to choose for a particular diagnostic task. As ratings we have chosen discrete values from 1 to 5 for the level of confidence (LOC) in a positive diagnosis. The meaning of these values was as follows: (1) definitely or almost definitely negative, (2) probably negative, (3) possibly positive, (4) probably positive, and (5) definitely or almost definitely positive. With this choice the value for the LOC is similar to the standard BI_RAD rating in mammographic screening.

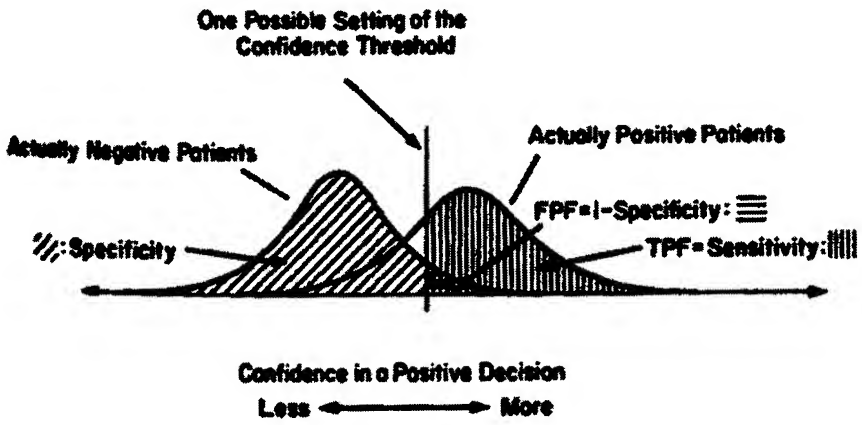



**Figure 12: Schematic example of the model that underlies ROC analysis. The bell-shaped curves represent probability density distributions of a radiologist's confidence in a positive diagnosis. A confidence threshold, represented by a vertical line, separates "positive" decisions from "negative" decisions (This figure was reprinted from[20]).**

To generate the ROC curve from discrete data, it is required to make assumptions about the functional form of the curve. The "binormal" model has been widely used in medical imaging. This model includes two adjustable parameters, and it is assumed that each conventional ROC curve has the same functional form as that implied by two "normal" (i.e., Gaussian) decision variable distributions with generally different means and standard deviations [21], [22]. It has the property that all possible ROC curves are transformed into straight lines, if they are plotted on "normal-deviate" axes [21], [22]. In effect, a "normal-deviate" graph stretches the unit square of the conventional ROC plot into an entire plane in a way such that the center of the unit square becomes the origin of the normal-deviate graph and the distance between any two points in the unit square is magnified increasingly as the points approach the border of the square.

The two adjustable parameters of the binormal ROC curve can be taken to be the y-intercept and the slope of the straight line that represents the ROC curve, when it is plotted on normal-deviate axes. These two parameters, denoted as "a' and "b", can be interpreted as an effective pair of underlying Gaussian distributions a s the distance between the means of the two distributions and the standard deviation of the actually negative distribution, respectively with both expressed in units of the standard deviation of the actually positive distribution [20]. With the binormal model a maximum-likelihood parameter estimation scheme is then used to generate an ROC curve that best represents the data.

If two different diagnostic systems are to be evaluated, the statistical difference of an apparent difference between measured ROC curves is of interest. For a detailed review of testing differences between ROC curves, the reader is referred to [23] and [24].

## 5. Results from ROC Analysis and Discussion

Meaningful ROC analysis was possible, since the "ground truth" for each case was provided by the mammography database. An initial analysis of the data counted the number of false-positives and true-positives in each group of cases. To consider a lesion as being diagnosed as malignant or benign, the LOC value was thresholded [20]. This threshold influences the shape of the ROC curve and its interpretation. In general, any enhancement protocol should increase sensitivity, i.e. fraction of true-positives (TPF), without decreasing specificity, i.e. essentially without increasing the fraction of false-positives (FPF) [25].

If the threshold for the level of confidence was chosen to be 3, meaning that lesions with a LOC greater or equal 3 were considered as malignant, then the average TPF was found to be 0.667 with enhancement, and TPF = 0.569 without enhancement. This observed increase in sensitivity is encouraging, though it was accompanied by a slight increase in the fraction of false-positives (0.222 compared to 0.178). The latter is not too surprising, *since the applied enhancement protocol only used dyadic spline wavelets* with the non-linear sigmoidal enhancement function, which is not be the optimal choice for all types of lesions. We believe that dyadic

wavelet expansions are best used to enhance microcalcifications. If the analysis of the data only focuses on microcalcifications, then we observed TPF = 0.417 with enhancement compared to TPF = 0.222 without enhancement. No increase or decrease in FPF was noticed! The last finding reinforces the idea for future research to design specific enhancement protocols for each mammographic feature.
Table 4 summarizes initial results of the first ROC study using a single basis function.

| With Enhancement (all Types of Lesions) | | Without Enhancement (all Types of Lesions) | |
|---|---|---|---|
| TPF | FPF | TPF | FPF |
| 0.667 | 0.233 | 0.569 | 0.178 |
| | | | |
| | | | |
| With Enhancement (Micros only) | | Without Enhancement (Micros only) | |
| TPF | FPF | TPF | FPF |
| 0.417 | 0.0 | 0.222 | 0.0 |

**Table 4: Results of preliminary ROC study. TPF refers to the fraction of true-positives and FPF to the fraction of false-positives.**

A more thorough analysis of the data was undertaken by using the *ROCKIT* software developed by the group led by Charles Metz at the University of Chicago [26]. This software was written to analyze data from ROC studies and to generate corresponding ROC curves. More specifically, the purpose of *ROCKIT* is to calculate maximum-likelihood estimates of the parameters of a conventional "binormal" model for the input data, to calculate maximum-likelihood estimates of the parameters of a "bivariate binormal" model for data from two potentially correlated diagnostic tests and, thus, to estimate the binormal ROC curves implied by those data and their correlation; and to calculate the statistical significance of the difference between two ROC curve estimates using any one of three distinct statistical tests:

1. The **Bivariate Test**: A bivariate Chi-square test of the simultaneous differences between the "a" parameters and between the "b" parameters of the two ROC curves. (*Null hypothesis*: the data sets arose from the same binormal ROC curve.)
2. The **Area Test:** A univariate z-score test of the difference between the areas under the two ROC curves. (*Null hypothesis*: the data sets arose from binormal ROC curves with equal areas beneath them.)
3. The **TFP Test:** A univariate z-score test of the difference between the true-positive fractions (TPFs) on the two ROC curves at a selected false-positive fraction (FPF). (*Null hypothesis*: the data sets arose from binormal ROC curves having the same TPF at the selected FPF.)

Three types of input data are allowed for statistical testing of the differences between ROC curves:

1. Unpaired (uncorrelated) test results. The two "conditions" are applied to independent case samples — for example, from two different diagnostic tests performed on the different patients, from two different radiologists who make probability judgments concerning the presence of a specified disease in different images, etc.;

2. Fully paired (correlated) test results, in which data from both of two conditions are available for each case in a single case sample. The two "conditions" in each test-result pair could correspond, for example, to two different diagnostic tests performed on the same patient, to two different radiologists who make probability judgments concerning the presence of a specified disease in the same image, etc.; and

3. Partially-paired test results — for example, two different diagnostic tests performed on the same patient sample and on some additional patients who received only one of the diagnostic tests.

*ROCKIT* assumes that the population ROC curve for each condition plots as a straight line on "normal-deviate" axes, or equivalently, that the input data follow normal distributions after some unknown monotonic transformation [20]. ROC curves measured in a broad variety of fields demonstrate this "binormal" form [27], [28], and [29]. The assumption may be satisfied even when the raw data have multimodal and/or skewed distributions. All this information was taken from [26].

Using the *ROCKIT* software the analysis was first applied independently to the datasets for group1 and group 2 for each of the three radiologists. Unfortunately, this approach did not lead to the desired result of being able to compare the diagnostic performance for the two diagnostic systems (softcopy display with and without enhancement). The reason for that was that the analysis for, at least one group was not completed, since the data was found to be degenerate [25]. In this case, the result of the ROC analysis would be a straight line with a constant value for TPF, and, therefore the software aborts processing to avoid meaningless output. According to the authors of the software, a degenerate data distribution can be found, if the number of samples is too small or in datasets with many tied values [26].
Since the number of cases could not be increased after conducting the study, and in order to obtain more complete results, we decided to apply the analysis to the union of data from all three radiologists. We found this decision justified by the fact that all the three radiologists came from the same population with a similar level of experience. Thus, their performance should be similar under the same conditions, and the data might be treated as independent samples. Nevertheless, we are well aware that the resulting statistical significance of the results has to be interpreted very carefully. For future ROC studies it is planned to increase the number of cases and to encourage the radiologists to make use of the full range of possible ratings for their level of confidence, in order to avoid such problems during the analysis of data.
For the software group 1 (with enhancement) was set as condition 1 and group 2 (without enhancement) was considered condition 2. The analysis of the overall data was carried out in two different ways. First, the data was regarded as unpaired (uncorrelated), since group 1 and group 2 contained different cases corresponding to independent samples. This interpretation of the data might be the most accurate one and was given most attention. For comparison and due to the fact that each mammographer diagnosed the same cases in group 1 and 2, the data was also analyzed as paired (correlated) data. The latter approach might be less correct, but was included in the report for completion.

On the next pages the resulting ROC curves for data analyzed as unpaired (see Figure 13 and Figure 14) and as paired (see Figure 15) together with their corresponding values for FPF and TPF (see Table 5 and Table 6, respectively) are shown. Figure 13 and Figure 14 refer to the same data, Figure 13 shows both curves in one diagram, while Figure 14 presents the curves separately. After that the most important results of ROC analysis, the binormal parameters a, b, and the area under the ROC curve $A_z$ with their corresponding standard errors, 95% confidence intervals, and correlation of a and b are summarized for unpaired data in Table 7 and for paired data in Table 8. Note that the 95% confidence intervals are symmetric for the binormal parameters a and b, but asymmetric for the area index $A_z$.
The complete output of the software *ROCKIT* for these two types of analysis is included in the appendix. As mentioned before group 1 corresponds to condition 1 and was abbreviated WE (with enhancement), and group 2 corresponds to condition 2, denoted as WOE (without enhancement).

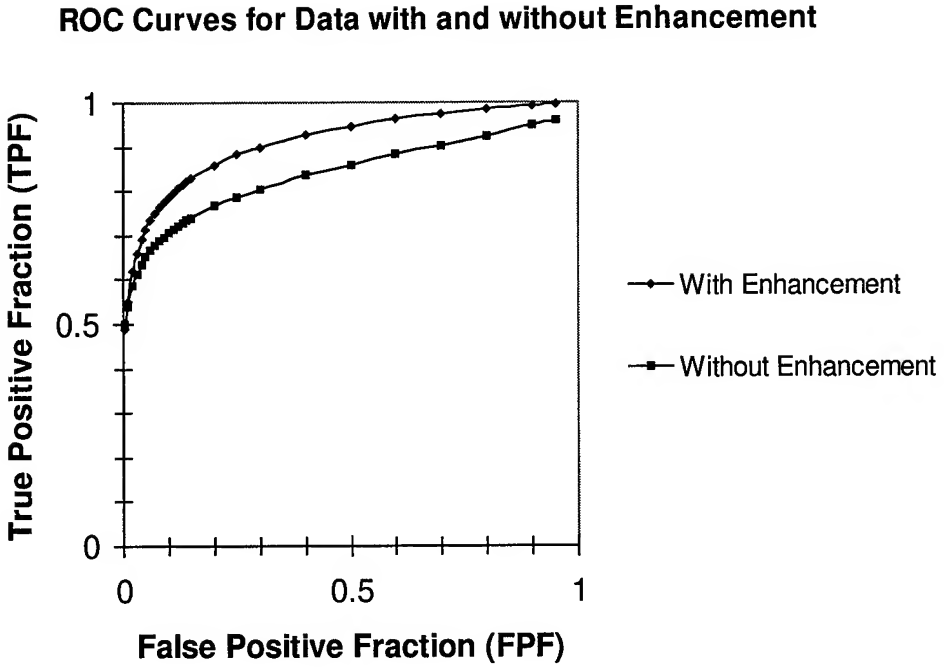



**Figure 13: ROC curves for data with condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as unpaired data (independent analysis).**

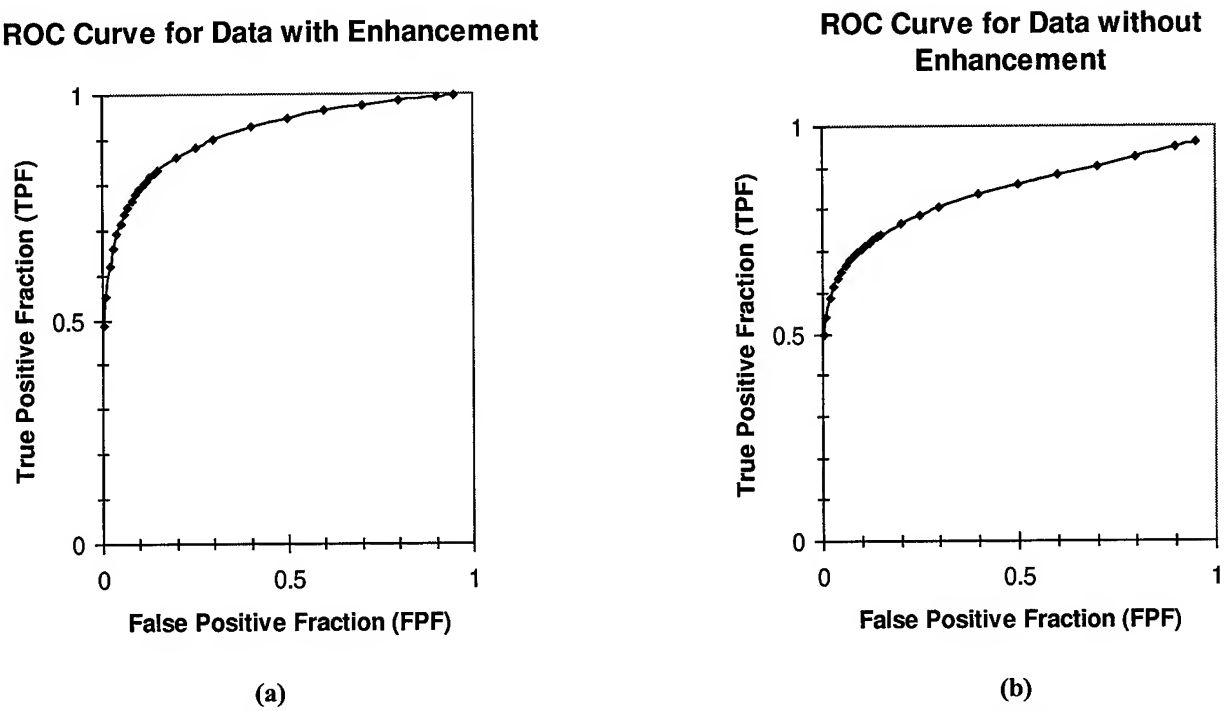



**Figure 14: ROC curves for data with (a) condition 1 (with enhancement) and (b) condition 2 (without enhancement) analyzed as unpaired data (independent analysis) in separate diagrams.**

| FPF | TPF |
|---|---|
| 0.005 | 0.4886 |
| 0.01 | 0.5521 |
| 0.02 | 0.6199 |
| 0.03 | 0.6612 |
| 0.04 | 0.6911 |
| 0.05 | 0.7145 |
| 0.06 | 0.7338 |
| 0.07 | 0.7501 |
| 0.08 | 0.7642 |
| 0.09 | 0.7767 |
| 0.1 | 0.7878 |
| 0.11 | 0.7979 |
| 0.12 | 0.8071 |
| 0.13 | 0.8155 |
| 0.14 | 0.8232 |
| 0.15 | 0.8304 |
| 0.2 | 0.86 |
| 0.25 | 0.8825 |
| 0.3 | 0.9003 |
| 0.4 | 0.9274 |
| 0.5 | 0.9472 |
| 0.6 | 0.9625 |
| 0.7 | 0.9746 |
| 0.8 | 0.9845 |
| 0.9 | 0.9926 |
| 0.95 | 0.9962 |

| FPF | TPF |
|---|---|
| 0.005 | 0.4989 |
| 0.01 | 0.5407 |
| 0.02 | 0.5859 |
| 0.03 | 0.614 |
| 0.04 | 0.6347 |
| 0.05 | 0.6514 |
| 0.06 | 0.6653 |
| 0.07 | 0.6773 |
| 0.08 | 0.6879 |
| 0.09 | 0.6974 |
| 0.1 | 0.7061 |
| 0.11 | 0.714 |
| 0.12 | 0.7213 |
| 0.13 | 0.7282 |
| 0.14 | 0.7346 |
| 0.15 | 0.7406 |
| 0.2 | 0.7665 |
| 0.25 | 0.7874 |
| 0.3 | 0.8053 |
| 0.4 | 0.8352 |
| 0.5 | 0.8602 |
| 0.6 | 0.8825 |
| 0.7 | 0.9035 |
| 0.8 | 0.9244 |
| 0.9 | 0.9475 |
| 0.95 | 0.9619 |

**Table 5: Values for false-positive fractions (FPF) and true-positive fractions (TPF) for condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as unpaired data (independent analysis).**

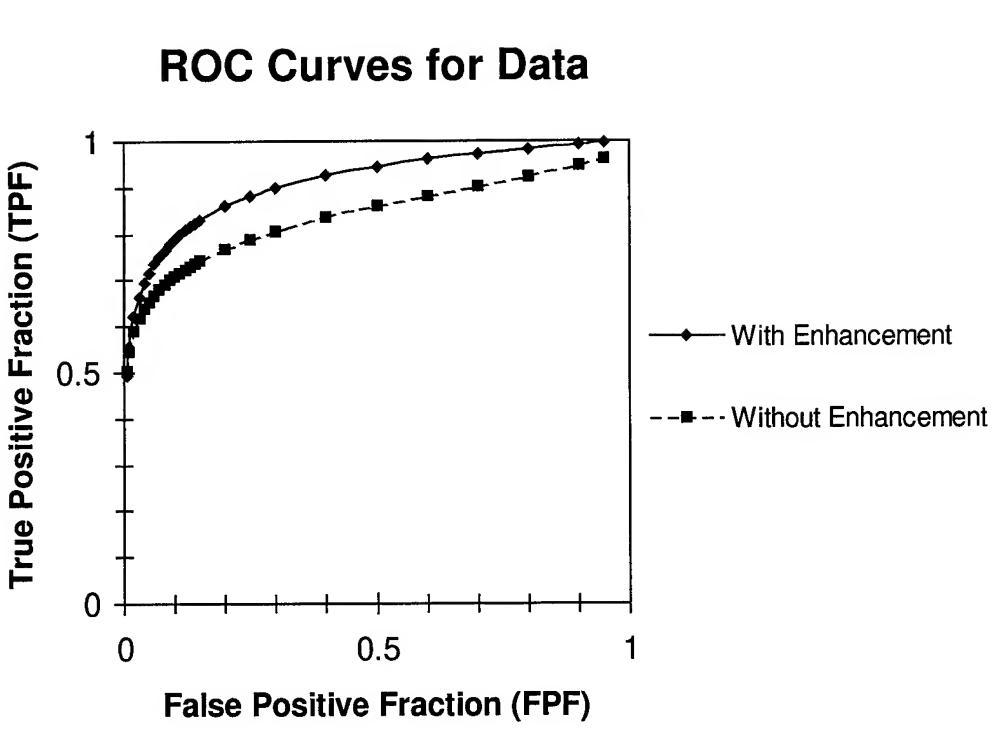



**Figure 15: ROC curves for data with condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as paired data.**

| FPF | TPF for Condition 1 | TPF for Condition 2 |
|---|---|---|
| 0.005 | 0.494 | 0.5036 |
| 0.01 | 0.5565 | 0.5451 |
| 0.02 | 0.6232 | 0.5898 |
| 0.03 | 0.6638 | 0.6176 |
| 0.04 | 0.6932 | 0.6381 |
| 0.05 | 0.7162 | 0.6545 |
| 0.06 | 0.7351 | 0.6683 |
| 0.07 | 0.7512 | 0.6801 |
| 0.08 | 0.7651 | 0.6906 |
| 0.09 | 0.7774 | 0.7 |
| 0.1 | 0.7883 | 0.7086 |
| 0.11 | 0.7982 | 0.7164 |
| 0.12 | 0.8073 | 0.7236 |
| 0.13 | 0.8155 | 0.7304 |
| 0.14 | 0.8232 | 0.7367 |
| 0.15 | 0.8303 | 0.7426 |
| 0.2 | 0.8595 | 0.7682 |
| 0.25 | 0.8817 | 0.7889 |
| 0.3 | 0.8994 | 0.8066 |
| 0.4 | 0.9263 | 0.8361 |
| 0.5 | 0.9461 | 0.8608 |
| 0.6 | 0.9614 | 0.8829 |
| 0.7 | 0.9737 | 0.9036 |
| 0.8 | 0.9838 | 0.9244 |
| 0.9 | 0.9922 | 0.9472 |
| 0.95 | 0.9959 | 0.9617 |

**Table 6: Values for false-positive fractions (FPF) and true-positive fractions (TPF) for condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as paired data.**

| Condition 1(With Enhancement) | | | Condition 2 (Without Enhancement) | | |
|---|---|---|---|---|---|
| Binormal Parameter a | Binormal Parameter b | Area under ROC Curve $A_z$ | Binormal Parameter a | Binormal Parameter b | Area under ROC Curve $A_z$ |
| 1.6183 | 0.6393 | 0.9136 | 1.0813 | 0.4208 | 0.8405 |
| | | | | | |
| | | | | | |
| Standard Error a | Standard Error b | Standard Error $A_z$ | Standard Error a | Standard Error b | Standard Error $A_z$ |
| 0.3162 | 0.2093 | 0.0325 | 0.2329 | 0.1307 | 0.0475 |
| | | | | | |
| | | | | | |
| 95% Confidence Interval for a | 95% Confidence Interval for b | 95% Confidence Interval for $A_z$ | 95% Confidence Interval for a | 95% Confidence Interval for b | 95% Confidence Interval for $A_z$ |
| (0.9986, 2.2381) | (0.2291, 1.0495) | (0.8312, 0.9615) | (0.6247, 1.5379) | (0.1647, 0.6770) | (0.7301, 0.9162) |
| | | | | | |
| | | | | | |
| | Correlation(a, b) | | | Correlation(a, b) | |
| | 0.6544 | | | 0.4989 | |

**Table 7: Binormal parameters a, b, area under ROC curve $A_z$ with their corresponding standard errors, 95% confidence intervals, and correlation(a, b) for condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as unpaired data (independent analysis).**

| Condition 1(With Enhancement) | | | Condition 2 (Without Enhancement) | | |
|---|---|---|---|---|---|
| Binormal Parameter a | Binormal Parameter b | Area under ROC Curve $A_z$ | Binormal Parameter a | Binormal Parameter b | Area under ROC Curve $A_z$ |
| 1.6084 | 0.6302 | 0.9132 | 1.0839 | 0.4172 | 0.8414 |
| | | | | | |
| | | | | | |
| Standard Error a | Standard Error b | Standard Error $A_z$ | Standard Error a | Standard Error b | Standard Error $A_z$ |
| 0.3137 | 0.2072 | 0.0327 | 0.233 | 0.1302 | 0.0474 |
| | | | | | |
| | | | | | |
| 95% Confidence Interval for a | 95% Confidence Interval for b | 95% Confidence Interval for $A_z$ | 95% Confidence Interval for a | 95% Confidence Interval for b | 95% Confidence Interval for $A_z$ |
| (0.9936, 2.2232) | (0.2240, 1.0363) | (0.8304, 0.9613) | (0.6272, 1.5407)) | (0.1620, 0.6724) | (0.7311, 0.9169) |
| | | | | | |
| | | | | | |
| | Correlation(a, b) | | | Correlation(a, b) | |
| | 0.6506 | | | 0.4995 | |

**Correlation of $A_z$ for condition 1 and $A_z$ for condition 2: -0.0922**

**Table 8: Binormal parameters a, b, area under ROC curve $A_z$ with their corresponding standard errors, 95% confidence intervals, and correlation (a, b) for condition 1 (with enhancement) and condition 2 (without enhancement) analyzed as paired data.**

As seen from both types of analysis, the values for the area under the ROC curve $A_z$ were larger for condition 1 (with enhancement) than they were for condition 2 (without enhancement). In all cases the standard error for $A_z$ was between 0.03 and 0.05, which was rather small. Though the 95% confidence intervals for $A_z$ overlapped, there was a clear tendency that diagnostic performance improved with enhancement in comparison with diagnosis without enhancement. All ROC curves lay high in the unit square of FPF and TPF, which

corresponded to accurate diagnostic performances in general, but the curves for condition 1 were positioned slightly higher (see Figure 13 and Figure 15). In general, results from data analyzed as unpaired and as paired were very similar. The small value of –0.0922 for the correlation of $A_z$ for condition 1 and condition 2 rather confirmed our suggestion that the data of the two conditions was unpaired.

The observed increase of the summary index $A_z$ within statistical errors encourages us to further pursue the application of enhancement protocols for mammographic screening. We are aware of the fact that there always are inherent sources of variability in the index $A_z$, such as a "case-sample" component due to random variations in the difficulty of the cases included in an ROC experiment, a "between-reader" component due to random variations in the skills of the observers participating in the experiment, and a "within-reader" component associated with each reader's inability to reproduce her/his diagnosis of every case on repeated readings [20]. In addition, we were not able to analyze the data for each radiologist separately due to data degeneracy as mentioned above. The latter has diminished the statistical significance of our results obtained from the analysis of all data combined, since not all samples were completely independent.

Hence, for future ROC studies we plan to increase the number of cases to avoid degenerate datasets for the analysis and to increase the statistical power of the experiment.

Aside from statistical considerations and the cautious interpretation of the results of this study we know that our prototype test bed software tool should be further optimized. To improve the enhancement protocol the idea is to develop feature specific enhancement protocols with different bases and associated non-linear functions for each distinct mammographic feature, such as microcalcifications, masses, and spicular lesions. The enhancement protocol used for this experiment, dyadic Spline wavelets with non-linear sigmoidal function, was suggested to work best for microcalcifications according to our previous work with multiscale expansions [16], [5]. The results of this first ROC experiment confirmed our expectations.

## D. Future Directions

As stated in previously above, one of our remaining goals is to further optimize the test bed software tool. One aspect of this is to achieve real-time processing for the reconstruction of a selected region of interest directly from multiscale coefficients. Our existing code can be sped up through the use of different types of filters, e.g. filter banks for biorthogonal wavelets, where computational operations for image reconstruction are reduced to fast multiplications.

Likewise, the choice of enhancement protocols will be expanded to a menu of feature specific enhancement algorithms tailored for each mammographic feature, such as microcalcifications, masses, and spicular lesions. A range of optimal choices for enhancement parameters to modify the corresponding enhancement functions will be investigated, possibly in response to an R-01, NIH program announcement in the area of digital mammography. Our "dream" is to present a clinical interface, where specific enhancement protocols can be selected by a physician by only pushing a button. We envision that through such a clinical interface the diagnostic performance of radiologists in screening can be substantially complemented and improved, both in terms of cost and quality.

Finally, more extensive ROC studies are planned to further evaluate the benefits of contrast enhancement through multiscale expansions for digitized and digital mammograms.

Some of these ideas have been recently proposed to the National Institute of Health (NIH) and the US Army Breast Cancer Research Program, and we hope to be able to continue this work with their support.

# Conclusions

In the paragraphs below, we summarize the results and progress made during the final year of the project. We identify the completion of tasks in *Phase IV* and *Phase V* investigated during this period in the Statement of Work revised in July 1997.

The first receiver operating characteristics (ROC) study to evaluate the benefits of contrast enhancement via overcomplete multiscale expansions of mammograms has been successfully completed. It was carried out in collaboration with radiologists and medical physicists at Columbia Presbyterian Medical Center of Columbia University.

In continuation of our previous work in digital mammography, an enhancement protocol using a dyadic Spline wavelet as the basis for multiscale expansion and an associated non-linear sigmoidal enhancement function was designed. Each digital mammogram was decomposed onto a multiscale basis to obtain coefficients at distinct subbands. Coefficients were modified by applying a non-linear sigmoidal function. Two parameters could be adjusted to change the enhancement. Image reconstruction from modified coefficients occurred in nearly real time through an interactive interface running on a "PACS style" digital mammography workstation.

To enable interactive feedback via high-speed processing during the ROC study, a graphical user interface (GUI) was designed. We called this interface the "test bed" software display tool. This display tool was implemented in Visual C++ 6.0 and allowed to load a complete case for a mammogram. All four tradional different views taken in mammography screening were displayed as downsampled images due to the large size of the digitized images. A selected view was connected to a viewport displaying a region of interest (ROI) at original resolution. The user could adjust the size of the square viewport.

The enhancement protocol was applied to the selected ROI for contrast enhancement of suspicious areas. Thus, the wavelet enhanced images provided a means of computer-aided diagnosis to the radiologist. Processing was limited to the ROI to achieve high speed for image reconstruction and to provide local enhancement for specific lesions. Multiple ROI's could be selected, processed, and the results saved.

In addition, to visualize raw data of digitized mammograms at the highest possible contrast and spatial resolutions, 16-Bit BARCO/Metheus framebuffers together with a dual headed high-resolution MegaScan grayscale monitor were utilized in the hardware setup. As formal members of the BARCO/Metheus Software Developer's Program (please see attached letter from company) we incorporated specialized software function calls to directly access the video framebuffer for fast image display and update.

To quantify the performance of our multiscale based processing technique in terms of overall sensitivity and specificity, an ROC study was designed and conducted with three radiologists from Columbia Presbyterian Medical Center specialized in mammography. Each mammographer diagnosed 60 cases of mammograms in two groups of 30 cases each according to the standard BI_RAD scale and a level of confidence (LOC) rating. The LOC values were in the range from 1 to 5 with 5 meaning the highest confidence in positive diagnosis (cancer) and 1 meaning the highest confidence in a negative diagnosis (normal). The usage of our enhancement algorithm was permitted to support diagnosis for the first group, but was not included in diagnosing the second group. Each group corresponded to a different condition of a distinct diagnostic system. Condition 1 was considered as softcopy display with enhancement, whereas condition 2 only corresponded to softcopy display. The study focused on dense mammograms of density 3 or 4 most difficult to asses for a physician. All cases were carefully selected from the national mammography database of digitized radiographs from the University of South Florida under the guidance of Dr. Suzanne Smith, Director of Breast Imaging at Columbia Presbyterian Hospital. We purchased the entire set of nearly 3000 cases from this national database. Additional resources, which were available to our group included a set of 300 cases of digital mammograms provided by LORAD/Bennett and access to their full-field digital mammography system, installed in our mammography center. The results of the ROC study were analyzed with the *ROCKIT* software provided by courtesy of Professor Charles Metz, Department of Radiology, University of Chicago [26]. Conventional ROC curves were generated and important statistical parameters determined. The area under the ROC curve $A_z$ was used as

summary index to quantify overall specificity and sensitivity of the two diagnostic systems [20]. Unfortunately, it was not possible to analyze datasets for each of three mammographers separately due to data degeneracy. Nevertheless, analyzing all the data together yielded a slight increase in the area $A_z$ for diagnosis with enhancement compared to diagnosis without. This result encourages us to further investigate the application of multiscale methods for contrast enhancement of mammograms, though we are also aware of the limited statistical significance of the obtained result. More extensive ROC studies with a larger number of cases are planned to further evaluate the benefits of our processing techniques.

Aside from statistical results we received very positive feedback from the participating radiologists, who expressed great interest in using the test bed software display tool and acknowledged a marked improvement in image quality, when enhancement was applied.

In summary, all of the proposed tasks described in *Phases IV* and *V* of this project have been successfully completed. The current enhancement protocol is best for the detection/enhancement of microcalcifications, and, as stated in the body of this report, we have started to apply the brushlet functions to mammograms with spicular lesions. Moreover, the subsection of Work in Progress under Section D in this report, suggested some possible new directions to be spun-off by this pioneering project. We hope that these efforts will be continued by us or other researchers, through NIH sponsor support in the near future.

# References


[1] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE*, vol. 84, pp. 626-638, 1996.

[2] M. Holschneider and R. Kronland-Martinet, "A real-time algorithm for signal analysis with the help of the wavelet transform," presented at Wavelets: Time-frequency Methods and Phase Space, Berlin, Germany, 1990.

[3] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D.J. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, pp. 587-607, 1992.

[4] S. D. Marco and J. Weiss, "M-band wavepacket-based transient signal detector using a translation-invariant wavelet," *Optical Engineering*, vol. 33, pp. 2175-2182, 1994.

[5] A. F. Laine, J. Fan, and W. Yang, "Wavelets for contrast enhancement of digital mammography," *IEEE Engineering in Medicine and Biology Society Magazine*, vol. 14, pp. 536-550, 1995.

[6] A. F. Laine, J. Fan, and S. Schuler, "A framework for contrast enhancement by dyadic wavelet analysis," in *Digital Mammography*, S. M. A. A. G. Gale, D. R. Dance, and A.Y. Cairns, Ed. Amsterdam, The Netherlands: Elsevier, 1994, pp. 91-100.

[7] C.-M. Chang and A. F. Laine, "Enhancement of mammograms from oriented information," Proceedings of the IEEE International Conference on Image Processing, Vol. III, pp. 524-527, Santa Barbara, CA, 1997.

[8] I. Koren and A. Laine, "A discrete dyadic wavelet transform for multidimensional feature analysis," in *Time frequency and wavelets in biomedical signal processing, IEEE press series in Biomedical Engineering*, M. Akay, Ed. New York: IEEE Press, 1998, pp. 425-448.

[9] P. G. Tahoces, J. Correa, M. Souto, and C. Gonzalez and, "Enhancement of chest and breast radiographs by automatic spatial filtering," *IEEE Transactions on Medical Imaging*, vol. 10, pp. 330-335, 1991.

[10] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Siam, 1992.

[11] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 710-732, 1992.

[12] E. Angelini, A. Laine, S. Takuma, and S. Homa, "Directional representations of 4D echocardiography for temporal quantification of LV volumes," submitted to Medical Imaging and Computer-Assisted Intervention - MICCAI'99, London, England, 1999.

[13] F. Meyer and R. R. Coifman, "Brushlets: A tool for directional image analysis and image compression," *Applied and computational harmonic analysis*, vol. 4, pp. 147-187, 1997.

[14] D. Chen, C.-M. Chang, and A. Laine, "Detection and enhancement of small masses via precision multiscale analysis," Proceedings of the Third Asian Conference on Computer Vision, vol. 1, pp. 192-199. Hong Kong, PRC, 1998.

[15] I. Koren, A. Laine, F. Taylor, and M. Lewis, "Interactive wavelet processing and techniques applied to digital mammography," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, pp. 1415-1418, Atlanta, GA, 1996.

[16] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 725-740, 1994.

[17] W. B. Richardson, Jr., "Nonlinear filtering and multiscale texture discrimination for mammograms," presented at Mathematical Methods in Medical Imaging, San Diego, CA, 1992.

[18] D. L. Donoho and I. M. Johnstone, "Threshold selection for wavelet shrinkage of noisy data," presented at 16th Annual Int. Conference of the IEEE Engineering in Medicine and Biology Society, 1994.

[19] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425-455, 1994.

[20] C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiology*, vol. 21, pp. 720-733, 1986.

[21] D. M. Green and J. A. Swets, *Signal detection theory and psychophysics*. New York: Wiley, 1966.

[22] J. A. Swets, "ROC analysis applied to the evaluation of medical imaging techniques," *Investigative Radiology*, vol. 14, pp. 109-121, 1979.

[23] B. J. Mc Neil and J. A. Hanley, "Statistical approaches to the analysis of receiver operating characteristic (ROC) curves," *Medical Decision Making*, vol. 4, pp. 137-150, 1984.

[24] J. A. Swets and R. M. Pickett, *Evaluation of diagnostic systems: Methods from signal detection theory*. New York: Academic Press, 1982.

[25] C. E. Metz, "Practical issues of experimental design and data analysis in radiological ROC studies," *Investigative Radiology*, vol. 24, pp. 234-245, 1989.

[26] C. E. Metz, B. A. Herman, and C. A. Roe, "Statistical comparison of two ROC curve estimates obtained from partially-paired datasets," *Medical Decision Making*, vol. 18, pp. 110-121, 1998.

[27] J. A. Swets, "Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance," *Psych Bull*, vol. 99, pp. 181-198, 1986.

[28] J. A. Hanley, "The robustness of the "binormal" assumptions used in fitting ROC curves," *Medical Decision Making*, vol. 8, pp. 197-203, 1988.

[29] K. O. Hajian-Tilaki, J. A. Hanley, L. Joseph, and J-P. Collet, "A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests," *Medical Decision Making*, vol. 17, pp. 94-107, 1997.

# Appendices

## Appendix I

As required of the final report, we have provided a complete list of publications and abstracts presented pertaining to results and studies of this project and a listing of all paid personnel affiliated with this grant.

**Chronological Summary of Published Works**

**Publications and Abstracts**


[1] A. F. Laine, J. Fan, and S. Schuler, "A framework for contrast enhancement by dyadic wavelet analysis," in *Digital Mammography*, S. M. A. A. G. Gale, D. R. Dance, and A.Y. Cairns, Ed. Amsterdam, The Netherlands: Elsevier, 1994, pp. 91-100.

[2] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 725-740, 1994.

[3] A. F. Laine, "A wavelet based mammographic system," presented at IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 5, pp. 21-24, Adelaide, Australia, 1994.

[4] A. F. Laine, J. Fan, and S. Schuler, "Contrast enhancement by dyadic wavelet analysis," , vol. 1, pp. 10a-11a, 1994.

[5] A. F. Laine, J. Fan, and S. Schuler, "Discrete dyadic wavelets for contrast enhancement," *Proceedings of SPIE - the International Society for Optical Engineering*, Vol. 2303, pp. 456-460, USA, 1994.

[6] A. Laine, *Wavelet Theory and Applications*. Boston, MA: Kluwer Academic Publishers, 1995.

[7] A. F. Laine, J. Fan, and W. Yang, "Wavelets for contrast enhancement of digital mammography,," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, pp. 536-550, 1995.

[8] A. Laine, W. Huda, B. G. Steinbach, and J. C. Honeyman, "Mammographic image processing using wavelet processing techniques," *European Radiology*, vol. 5, pp. 518-523, 1995.

[9] A. Laine, I. Koren, W. Yang, and F. Taylor, "A steerable dyadic wavelet transform and interval wavelets for enhancement of digital mammography," , vol. 2491, pp. 736-749, 1995.

[10] Y. Xing, W. Huda, A. Laine, J. Fan, and B. Steinbach, "Comparison of a dyadic wavelet image enhancement algorithm with unsharp masking and median filtering," presented at Medical Imaging: Image Processing, Proceedings of the SPIE, vol. 2434, pp. 718-729, San Diego, CA, 1995.

[11] A. F. Laine and C.-M. Chang, "De-noising via wavelet transforms using steerable filters," , vol. 3, pp. 1956-1959, 1995.

[12] S. Schuler, B. Steinbach, and A. F. Laine, "Computer-aided diagnosis of digital mammography and wavelet representations," presented at Radiological Society of North American (RSNA) 81st Scientific Assembly and Annual Meeting, Chicago, Illinois, 1995.

[13] J. Fan and A. F. Laine, "Multiscale contrast enhancement and denoising in digital radiographs," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. Boca Raton, FL: CRC Press, 1996, pp. 163-189.

[14] A. F. Laine, W. Huda, D. Chen, and J. Harris, "Segmentation of masses using continuous scale representations," presented at Proceedings of the Third International Workshop on Mammography, pp. 447-450, Chicago, I.L., 1996.

[15] A. F. Laine and J. Fan, "Frame representations for texture segmentation," *IEEE Transactions on Image Processing*, vol. 5, pp. 771-783, 1996.

[16] I. Koren, A. Laine, F. Taylor, and M. Lewis, "Interactive wavelet processing and techniques applied to digital mammography," presented at Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pp. 1415-1418, Atlanta, GA, 1996.

[17] A. Laine, I. Koren, W. Huda, B. Steinbach, and J. Honeyman, "A tutorial of Wavelet Processing Techniques for Image Analysis," presented at Radiological Society of North America (RSNA) 82nd Scientific Assembly and Annual Meeting, pp. 557, Chicago, Illinois, 1996.

[18] A. Laine and X. Zong, "A multiscale sub-octave wavelet transform for denoising and enhancement," presented at Wavelet Applications in signal and Image Processing IV, Proceedings of SPIE, Vol. 2825, pp. 238-249, Denver, CO, 1996.

[19] X. Zong, A. Laine, E. Geiser, and D. Wilson, "Denoising and contrast enhancement via wavelet shrinkage and nonlinear adaptive gain," presented at Wavelet Applications III, Proceedings of SPIE, Vol. 2762, pp. 566-574, Orlando, FL, 1996.

[20] D. Chen, J. Harris, and A. Laine, "Automatic scale detection," presented at Visual Communications and Image Processing, Proceedings of SPIE, vol. 2727, pp. 960-972, Orlando, FL, 1996.

[21] B. Kalman, W. Reinus, S. Kwasny, A. Laine, and L. Kotner, "Prescreening Entire Mammograms for Masses with Artificial Neural Networks: Preliminary Results," *Academic Radiology*, vol. 4, pp. 405-414, 1997.

[22] C.-M. Chang and A. F. Laine, "Enhancement of mammograms from oriented information," presented at Proceedings of the IEEE International Conference on Image Processing, Vol. III, pp. 524-527, Santa Barbara, CA, 1997.

[23] X. Zong, A. Meyer-Baese, and A. Laine, "Multiscale Segmentation Through A Radial Basis Neural Network," presented at International Conference on Image Processing, vol. III, pp. 400-403, Santa Barbara, CA, 1997.

[24] A. Laine, M. Shim, I. Koren, W. Huda, and B. Steinbach, "Multiscale Processing Techniques for the Enhancement of Digital Radiographs," presented at Radiological Society of North America (RSNA) 83rd Scientific Assembly and Annual Meeting, Chicago, Illinois, 1997.

[25] S. Schuler and A. Laine, "Hexagonal QMF banks and wavelets," in *Time frequency and wavelets in biomedical signal processing, IEEE press series in Biomedical Engineering*, M. Akay, Ed., 1998, pp. 449-472.

[26] I. Koren and A. Laine, "A discrete dyadic wavelet transform for multidimensional feature analysis," in *Time frequency and wavelets in biomedical signal processing, IEEE press series in Biomedical Engineering*, M. Akay, Ed. New York: IEEE Press, 1998, pp. 425-448.

[27] A. F. L. Xuli Zong, Edward A. Geiser, "Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 532-540, 1998.

[28] M. Shim and A. Laine, "Overcomplete Lifted Representations for Multiscale Feature Analysis," presented at International Conference on Image Processing, vol. 2, pp. 242-246, Chicago, IL, 1998.

[29] I. Koren, A. Laine, and F. Taylor, "Enhancement via Fusion of Mammographic Features," presented at IEEE International Conference on Image Processing, vol. 2, pp. 722-726, Chicago, IL, 1998.

[30] C. Chang, D. Chen, and A. Laine, "Detection of Subtle Masses in Mammogrphy via Multiscale Analysis at Arbitrary Sceles," presented at Third Asian Conference on Computer Vision, vol. 1, pp. 192-199, Hong Kong, 1998.

[31] A. Laine, *Emerging Technologies in Biomedical Engineering: Wavelets and Medical Image Processing.*: IEEE Press and Engineering in Medicine and Biology Society (EMBS), expected in summer 1999.

[32] A. Laine, W. Huda, and R. Mekle, "Enhancement of Digital Radiographs via Multiscale Representations," in *Handbook of Medical Image Processing*, R. Rangayyan, Ed.: Academic Press, 1999.

[33] C. Chang and A. Laine, "Coherence of Multiscale Features for Contrast Enhancement of Dense Radiographs," *IEEE Transactions on Information Technology in Biomedicine*, vol. 3, pp. 32-46, 1999.

[34] D. Ioannou, W. Huda, and A. Laine, "Circle recognition through a 2D hough transform and radius histogramming," *Image and Vision Computing*, vol. 17, pp. 27-36, 1999.

[35] A. Laine, D. Chen, J. Harris, and W. Huda, "Detection of Subtle Masses in Mammography via Analysis at Arbitrary Scales," in *Image Processing Techniques for Tumor Detection*, R. Strickland, Ed.: Marcel Dekker, Inc., to appear.

[36] I. Koren and A. Laine, "Enhancement of Mammography via Fusion of Multiscale Features," in *First International Workshop on Computer aided Diagnosis*, M. Doi, Ed. Amsterdam: Elsevier Science, in press.

**Lectures and Invited talks**

[1] International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), November 1-2, 1994, Baltimore, MD, invited speaker.

[2] International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Workshop on Wavelets in Medicine, September, 1995, Montreal, Canada, invited speaker.

[3] Argonne National Laboratory, Mathematics and Computer Science Division, Workshop on Large Scale Medical Imaging, August, 1995, invited talk.

[4] National Institutes of Health (NIH), National Cancer Institute, Early Detection Division, Workshop on Double Reading Mammography, invited speaker on computer-assisted reading, Bethesda, MD, October 30-31, 1995.

[5] Technische Hochschule (Universitaet) Darmstadt, Darmstadt, Germany, July 1997, invited lecturer. MeVis, Center for Diagnostic Systems and Visualization, Bremen, Germany, July 1997, invited speaker.

[6] U.S. Army Medical Research and Material Command, Breast Cancer Research Program: An Era of Hope, Washington, DC, October 1997, poster presentation.

[7] National Science Foundation, NSF Regional Conference on Harmonic Analysis, invited speaker, Orlando, FL, May 1998.

[8] IEEE Engineering in Medicine and Biology Society, 20$^{th}$ Annual International Conference, Instructor for Tutorial Workshop on Wavelets in Medicine and Biology, Hong Kong, October 28, 1998.

**List of Personnel**

| | |
|---|---|
| Andrew Laine, D.Sc., | Principal Investigator |
| Walter Huda, Ph.D., | Medical Physicist |
| Barbara Steinbach, M.D., | Mammographer |
| Janice Honeyman, Ph.D., | Computer Scientist |
| Susan Smith, M.D., | Mammographer |
| Edward Nickoloff, Ph.D. | Medical Physicist |
| Jian Fan, Ph.D., | Doctoral Degree Graduate Student |
| Sergio Schuler, Ph.D., | Doctoral Degree Graduate Student |
| Iztok Koren, Ph.D., | Doctoral Degree Graduate Student |
| Chun-Ming Chang, Ph.D., | Doctoral Degree Graduate Student |
| Ralf Mekle, M.S. | Doctoral Degree Graduate Student |

## Appendix II

In this appendix we include the complete output of the ROC software *ROCKIT* for the data analyzed as unpaired and as paired input.

## 1. Output for Data Analyzed as Unpaired Input

**Condition 1 (With Enhancement = WE)**

WE WOE

Date - 08-Jun-99
Time - 21:03:55

ROCKIT (Windows95 version 0.9.1 BETA):

Maximum Likelihood Estimation of a Binormal ROC Curve
from RATING Data

Original Categorical Response Data:

With category runs collapsed.

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 22 | 17 | 6 | 0 | 0 |
| Actually-Positive Cases | 2 | 7 | 15 | 16 | 5 |

Date - 08-Jun-99
Time - 21:03:55

ROCKIT (Windows95 version 0.9.1 BETA):
Enhancement on mammo, Pooled Data, with or without

Maximum Likelihood Estimation of the Parameters
a Single Binormal ROC Curve

Name of Input File being used: Pooled_ROC_Input.prn

Condition 1: WE

Total number of actually-negative cases = 45.
Total number of actually-positive cases = 45.

Data effectively collected in 5 categories.
Category 5 represents the strongest evidence of positivity.
(e.g., that the disease is present)

Response Data:

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 22 | 17 | 6 | 0 | 0 |
| Actually-Positive Cases | 2 | 7 | 15 | 16 | 5 |

Operating Points Corresponding to the Input Data:

FPF: .000 .000 .000 .133 .511 1.000
TPF: .000 .111 .467 .800 .956 1.000

---

Initial Estimates of the Binormal ROC Parameters:

---

a = 1.7949
b = 1.0319

z(k) = -.002 -1.005 -1.917 -2.936

Procedure Converges after 7 Iterations

==========

Final Estimates of the Binormal ROC Parameters

==========

Binormal Parameters and Area Under the Estimated ROC:

| | | |
|---|---|---|
| a | = | 1.6183 |
| b | = | .6393 |
| Area (Az) | = | .9136 |

1: z(k) = -.037 1.153 2.676 4.448

Estimated Standard Errors and Correlation of these Values:

| | | |
|---|---|---|
| Std. Err. (a) | = | .3162 |
| Std. Err. (b) | = | .2093 |
| Corr(a,b) | = | .6544 |
| Std. Err. (Az) | = | .0325 |

Symmetric 95% Confidence Intervals

| | |
|---|---|
| For a : | ( .9986, 2.2381) |
| For b : | ( .2291, 1.0495) |

Asymmetric 95% Confidence Interval

| | |
|---|---|
| For Az: | ( .8312, .9615) |

Variance-Covariance Matrix:

==========

| | a | b | z( 1) | z( 2) | z( 3) | z( 4) |
|---|---|---|---|---|---|---|
| a | .1000 | | | | | |
| b | .0433 | .0438 | | | | |
| z( 1) | .0206 | .0074 | .0347 | | | |
| z( 2) | .0177 | -.0101 | .0174 | .0527 | | |
| z( 3) | -.0429 | -.1047 | .0014 | .0684 | .3824 | |
| z( 4) | -.1539 | -.2314 | -.0193 | .0969 | .6667 | 1.4668 |

Correlation Matrix:

| | a | b | z( 1) | z( 2) | z( 3) | z( 4) |
|---|---|---|---|---|---|---|
| a | 1.0000 | | | | | |
| b | .6544 | 1.0000 | | | | |
| z( 1) | .3499 | .1906 | 1.0000 | | | |
| z( 2) | .2438 | -.2092 | .4066 | 1.0000 | | |
| z( 3) | -.2194 | -.8088 | .0125 | .4813 | 1.0000 | |
| z( 4) | .0000 | .0000 | .0000 | .0000 | .0000 | .0000 |

Estimated Binormal ROC curve, with Lower and Upper Bounds of the Asymmetric Point-wise 95% Confidence Interval for True-Positive Fraction at a Variety of False-Positive Fractions:

| FPF | TPF | (Lower Bound, Upper Bound) |
|---|---|---|
| .005 | .4886 | ( .2030 , .7804 ) |
| .010 | .5521 | ( .2773 , .8031 ) |
| .020 | .6199 | ( .3686 , .8279 ) |
| .030 | .6612 | ( .4290 , .8438 ) |
| .040 | .6911 | ( .4743 , .8559 ) |
| .050 | .7145 | ( .5104 , .8659 ) |
| .060 | .7338 | ( .5403 , .8744 ) |
| .070 | .7501 | ( .5656 , .8818 ) |
| .080 | .7642 | ( .5875 , .8885 ) |
| .090 | .7767 | ( .6067 , .8946 ) |
| .100 | .7878 | ( .6237 , .9002 ) |
| .110 | .7979 | ( .6389 , .9054 ) |
| .120 | .8071 | ( .6526 , .9102 ) |
| .130 | .8155 | ( .6650 , .9147 ) |
| .140 | .8232 | ( .6764 , .9189 ) |
| .150 | .8304 | ( .6868 , .9229 ) |
| .200 | .8600 | ( .7284 , .9398 ) |
| .250 | .8825 | ( .7585 , .9529 ) |
| .300 | .9003 | ( .7816 , .9632 ) |
| .400 | .9274 | ( .8157 , .9780 ) |
| .500 | .9472 | ( .8410 , .9874 ) |
| .600 | .9625 | ( .8617 , .9933 ) |
| .700 | .9746 | ( .8802 , .9968 ) |
| .800 | .9845 | ( .8982 , .9988 ) |
| .900 | .9926 | ( .9185 , .9997 ) |
| .950 | .9962 | ( .9322 , .9999 ) |

Estimates of Expected Operating Points on fitted ROC curve, with lower and upper bounds of asymmetric 95% confidence interval along the curve for those points:

| Expected operating point ( FPF , TPF ) | Lower bound ( FPF , TPF ) | Upper bound ( FPF , TPF ) |
|---|---|---|
| ( .0037, .4633) | (0.0000, .0537) | ( .3814, .9230) |
| ( .1245, .8109) | ( .0090, .5424) | ( .5235, .9511) |
| ( .5147, .9497) | ( .3397, .9122) | ( .6869, .9732) |
| ( .5000, .9472) | ( .3575, .9170) | ( .6425, .9680) |

Date - 08-Jun-99
Time - 21:03:55

Condition 2 (Without Enhancement = WOE)

Date - 08-Jun-99
Time - 21:04:30

ROCKIT (Windows95 version 0.9.1 BETA):

Maximum Likelihood Estimation of a Binormal ROC Curve
from RATING Data

Original Categorical Response Data:

With category runs collapsed.

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 18 | 21 | 6 | 0 | 0 |
| Actually-Positive Cases | 5 | 8 | 10 | 16 | 6 |

Date - 08-Jun-99
Time - 21:04:30

ROCKIT (Windows95 version 0.9.1 BETA):
Enhancement on mammo, Pooled Data, with or without

Maximum Likelihood Estimation of the Parameters
a Single Binormal ROC Curve

Name of Input File being used: Pooled_ROC_Input.prn

Condition 1: WOE

Total number of actually-negative cases = 45.
Total number of actually-positive cases = 45.

Data effectively collected in 5 categories.
Category 5 represents the strongest evidence of positivity.
(e.g., that the disease is present)

Response Data:

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 18 | 21 | 6 | 0 | 0 |
| Actually-Positive Cases | 5 | 8 | 10 | 16 | 6 |

Operating Points Corresponding to the Input Data:

FPF: .000 .000 .000 .133 .600 1.000
TPF: .000 .133 .489 .711 .889 1.000

---

Initial Estimates of the Binormal ROC Parameters:

---

a = 1.1537
b = .7188

z(k) = .217 -.994 -1.864 -3.163

Procedure Converges after 7 Iterations

==========

Final Estimates of the Binormal ROC Parameters

==========

Binormal Parameters and Area Under the Estimated ROC:

a = 1.0813
b = .4208
Area (Az) = .8405

1: z(k) = -.263 1.152 2.656 5.219

Estimated Standard Errors and Correlation of these Values:

Std. Err. (a) = .2329
Std. Err. (b) = .1307
Corr(a,b) = .4989
Std. Err. (Az) = .0475

Symmetric 95% Confidence Intervals
For a : ( .6247, 1.5379)

For b : ( .1647, .6770)

Asymmetric 95% Confidence Interval
For Az: ( .7301, .9162)

Variance-Covariance Matrix:
==========

| | a | b | z( 1) | z( 2) | z( 3) | z( 4) |
|---|---|---|---|---|---|---|
| a | .0543 | | | | | |
| b | .0152 | .0171 | | | | |
| z( 1) | .0130 | .0050 | .0356 | | | |
| z( 2) | .0108 | -.0066 | .0149 | .0533 | | |
| z( 3) | -.0072 | -.0583 | -.0003 | .0654 | .3663 | |
| z( 4) | -.2338 | -.4003 | -.0503 | .2036 | 1.3801 | 3.5527 |

Correlation Matrix:
==========

| | a | b | z( 1) | z( 2) | z( 3) | z( 4) |
|---|---|---|---|---|---|---|
| a | 1.0000 | | | | | |
| b | .4989 | 1.0000 | | | | |
| z( 1) | .2955 | .2032 | 1.0000 | | | |
| z( 2) | .2005 | -.2193 | .3413 | 1.0000 | | |
| z( 3) | -.0510 | -.7365 | -.0025 | .4681 | 1.0000 | |
| z( 4) | .0000 | .0000 | .0000 | .0000 | .0000 | .0000 |

Estimated Binormal ROC curve, with Lower and Upper Bounds of the Asymmetric Point-wise 95% Confidence Interval for True-Positive Fraction at a Variety of False-Positive Fractions:

| FPF | TPF | (Lower Bound, Upper Bound) |
|---|---|---|
| .005 | .4989 | ( .2780 , .7201 ) |
| .010 | .5407 | ( .3306 , .7398 ) |
| .020 | .5859 | ( .3902 , .7619 ) |
| .030 | .6140 | ( .4284 , .7763 ) |
| .040 | .6347 | ( .4567 , .7874 ) |
| .050 | .6514 | ( .4794 , .7966 ) |
| .060 | .6653 | ( .4984 , .8045 ) |
| .070 | .6773 | ( .5147 , .8115 ) |
| .080 | .6879 | ( .5290 , .8178 ) |
| .090 | .6974 | ( .5417 , .8236 ) |
| .100 | .7061 | ( .5532 , .8290 ) |
| .110 | .7140 | ( .5636 , .8340 ) |
| .120 | .7213 | ( .5732 , .8387 ) |
| .130 | .7282 | ( .5820 , .8432 ) |
| .140 | .7346 | ( .5902 , .8474 ) |
| .150 | .7406 | ( .5978 , .8514 ) |
| .200 | .7665 | ( .6298 , .8693 ) |
| .250 | .7874 | ( .6547 , .8844 ) |
| .300 | .8053 | ( .6752 , .8975 ) |
| .400 | .8352 | ( .7078 , .9197 ) |
| .500 | .8602 | ( .7339 , .9380 ) |
| .600 | .8825 | ( .7567 , .9535 ) |
| .700 | .9035 | ( .7780 , .9670 ) |
| .800 | .9244 | ( .7999 , .9788 ) |
| .900 | .9475 | ( .8259 , .9894 ) |
| .950 | .9619 | ( .8446 , .9944 ) |

Estimates of Expected Operating Points on fitted ROC curve, with lower and upper bounds of asymmetric 95% confidence interval along the curve for those points:

| Expected operating point ( FPF , TPF ) | Lower bound ( FPF , TPF ) | Upper bound ( FPF , TPF ) |
|---|---|---|
| ( .0040, .4856) | (0.0000, .0558) | ( .8506, .9355) |
| ( .1247, .7246) | ( .0097, .5388) | ( .5138, .8634) |
| ( .6039, .8834) | ( .4250, .8418) | ( .7629, .9166) |
| ( .5000, .8602) | ( .3558, .8227) | ( .6442, .8919) |

Date - 08-Jun-99
Time - 21:04:30

## 2. Output for Data Analyzed as Paired Input

Output for the analysis of data for both, condition 1 and 2, was written to one file in this case.

Condition 1 (With Enhancement = WE), Condition 2 (Without Enhancement = WOE) Analyzing file comparing datasets 1 & 2

Date - 08-Jun-99
Time - 20:57:23

ROCKIT (Windows95 version 0.9.1 BETA):

Maximum Likelihood Estimation of a Binormal ROC Curve
from RATING Data

for condition 1 : WE

Original Categorical Response Data:

With category runs collapsed.

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 22 | 17 | 6 | 0 | 0 |
| Actually-Positive Cases | 2 | 7 | 15 | 16 | 5 |

from RATING Data

for condition 2 : WOE

Original Categorical Response Data:

With category runs collapsed.

| Category | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Actually-Negative Cases | 18 | 21 | 6 | 0 | 0 |
| Actually-Positive Cases | 5 | 8 | 10 | 16 | 6 |

Date - 08-Jun-99
Time - 20:57:26

ROCKIT (Windows95 version 0.9.1 BETA):
Enhancement on mammo, Pooled Data, with or without

Maximum Likelihood Estimation of the Parameters
of the Bivariate Binormal Model for PAIRED Data
and
the Calculation of the Statistical Significance of
the Difference between Binormal ROC Curve Estimates.

Statistical Test to be Employed:
Area (Az) test

Name of Input File being used: Pooled_ROC_Input.prn

Condition 1: WE

Data effectively collected in 5 categories.
Category 5 represents the strongest evidence of positivity.
(e.g., that the disease is present)

Condition 2: WOE

Data effectively collected in 5 categories.
Category 5 represents the strongest evidence of positivity.
(e.g., that the disease is present)

Total number of correlated actually-negative cases = 45.
Total number of correlated actually-positive cases = 45.

Rating-Data Matrix for Actually-Negative cases:

| | Condition 1 Ratings | | | | | |
|---|---|---|---|---|---|---|
| Condition 2 Ratings | 1 | 2 | 3 | 4 | 5 | |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 5 | 1 | 0 | 0 | 0 | 6 |
| 2 | 5 | 11 | 5 | 0 | 0 | 21 |
| 1 | 12 | 5 | 1 | 0 | 0 | 18 |
| sum 1 | 22 | 17 | 6 | 0 | 0 | 45 |

Rating-Data Matrix for Actually-Positive cases:

| | Condition 1 Ratings | | | | | |
|---|---|---|---|---|---|---|
| Condition 2 Ratings | 1 | 2 | 3 | 4 | 5 | |
| 5 | 1 | 1 | 3 | 1 | 0 | 6 |
| 4 | 0 | 2 | 10 | 3 | 1 | 16 |
| 3 | 1 | 2 | 0 | 4 | 3 | 10 |
| 2 | 0 | 1 | 1 | 5 | 1 | 8 |
| 1 | 0 | 1 | 1 | 3 | 0 | 5 |
| sum 1 | 2 | 7 | 15 | 16 | 5 | 45 |

Operating Points Corresponding to the Input Data:

For Condition 1:
FPF: .000 .000 .000 .133 .511 1.000
TPF: .000 .111 .467 .800 .956 1.000

Operating Points Corresponding to the Input Data:

For Condition 2:
FPF: .000 .000 .000 .133 .600 1.000
TPF: .000 .133 .489 .711 .889 1.000

---

Initial Estimates of the Binormal ROC Parameters:

---

For Condition 1: WE

a = 1.7949
b = 1.0319

z(k) = -.002 -1.005 -1.917 -2.936

---

Initial Estimates of the Binormal ROC Parameters:

---

For Condition 2: WOE

a = 1.1537
b = .7188

z(k) = .217 -.994 -1.864 -3.163

Procedure Converges after 4 Iterations

==========

Final Estimates of the Binormal ROC Parameters
and the Inter-Condition Correlation Coefficients:

==========

| | Condition 1: WE | Condition 2: WOE |
|---|---|---|
| Binormal Parameters and Area Under the Estimated ROC : | | |
| a = | 1.6084 | 1.0839 |
| b = | .6302 | .4172 |
| Area (Az) = | .9132 | .8414 |

*** Wilcoxon area estimates are computed for continuous data only.

1: z(k) = -.036 1.151 2.693 4.510
2: z(k) = -.256 1.144 2.658 5.289

| | | |
|---|---|---|
| Estimated Standard Errors and Correlation of these Values: | | |
| Std. Err. (a) = | .3137 | .2330 |
| Std. Err. (b) = | .2072 | .1302 |
| Corr(a,b) = | .6506 | .4995 |
| Std. Err. (Az) = | .0327 | .0474 |

*** Wilcoxon area estimates are computed for continuous data only.

| | | |
|---|---|---|
| Symmetric 95% Confidence Intervals | | |
| For a : | ( .9936, 2.2232) | ( .6272, 1.5407) |
| For b : | ( .2240, 1.0363) | ( .1620, .6724) |

| | | |
|---|---|---|
| Asymmetric 95% Confidence Interval | | |
| For Az: | ( .8304, .9613) | ( .7311, .9169) |

Inter-Condition Decision Variable Correlation Estimates:
Effective Correlation of the Test Results Between Conditions:
For Actually-Negative Cases (Rn) = .1301

For Actually-Positive Cases (Rs) = -.2980

Estimated Standard Errors of the Inter-Condition
Correlation Coefficients:
Std. Error of Rn (for Actually-Negative Cases)= .1887
Std. Error of Rs (for Actually-Positive Cases)= .1511

Correlation of Area(1) and Area(2) = -.0922

Statistical Significance of the Difference between the Two
*CORRELATED* ROC Curve Estimates According to the Selected Test:
**********

The computed *CORRELATED* Area test statistic
has a value of = 1.1959

with corresponding two-tailed p-value = .2317
and corresponding one-tailed p-value = .1159.

Approximate 95% Confidence Interval for the Difference:
(-.0459, .1894)

Variance-Covariance Matrix:
==========

| | ax | bx | ay | by | rs | rn | zx( 1) | zx( 2) | zx( 3) | zx( 4) | zy( 1) | zy( 2) | zy( 3) | zy( 4) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ax | .0984 | | | | | | | | | | | | | |
| bx | .0423 | .0429 | | | | | | | | | | | | |
| ay | -.0041 | .0007 | .0543 | | | | | | | | | | | |
| by | .0006 | .0003 | .0152 | .0169 | | | | | | | | | | |
| rs | .0050 | .0020 | .0035 | .0014 | .0228 | | | | | | | | | |
| rn | .0001 | .0011 | 0.0000 | .0007 | 0.0000 | .0356 | | | | | | | | |
| zx( 1) | .0203 | .0073 | .0012 | 0.0000 | 0.0000 | .0003 | .0347 | | | | | | | |
| zx( 2) | .0173 | -.0100 | .0012 | -.0001 | 0.0000 | -.0018 | .0174 | .0527 | | | | | | |
| zx( 3) | -.0433 | -.1057 | .0012 | -.0002 | -.0001 | -.0044 | .0012 | .0689 | .3948 | | | | | |
| zx( 4) | -.1555 | -.2346 | .0013 | -.0002 | -.0004 | -.0075 | -.0201 | .0985 | .6931 | 1.5328 | | | | |
| zy( 1) | .0018 | 0.0000 | .0129 | .0050 | 0.0000 | .0005 | .0029 | .0028 | .0027 | .0026 | .0355 | | | |
| zy( 2) | .0018 | -.0001 | .0107 | -.0065 | 0.0000 | -.0017 | .0029 | .0031 | .0035 | .0039 | .0150 | .0529 | | |
| zy( 3) | .0018 | -.0003 | -.0074 | -.0584 | 0.0000 | -.0041 | .0028 | .0035 | .0043 | .0051 | -.0002 | .0648 | .3701 | |
| zy( 4) | .0022 | -.0003 | -.0824 | -.1717 | -.0008 | -.0084 | .0028 | .0041 | .0055 | .0056 | -.0317 | .1065 | .7299 | 2.1640 |

ax bx ay by rs rn zx( 1) zx( 2) zx( 3) zx( 4) zy( 1) zy( 2) zy( 3) zy( 4)

Correlation Matrix:
==========

| | ax | bx | ay | by | rs | rn | zx( 1) | zx( 2) | zx( 3) | zx( 4) | zy( 1) | zy( 2) | zy( 3) | zy( 4) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ax | 1.0000 | | | | | | | | | | | | | |
| bx | .6506 | 1.0000 | | | | | | | | | | | | |
| ay | -.0565 | .0139 | 1.0000 | | | | | | | | | | | |
| by | .0155 | .0106 | .4995 | 1.0000 | | | | | | | | | | |
| rs | .1061 | .0641 | .0999 | .0706 | 1.0000 | | | | | | | | | |
| rn | .0018 | .0273 | .0011 | .0272 | 0.0000 | 1.0000 | | | | | | | | |
| zx( 1) | .3481 | .1900 | .0278 | .0004 | -.0001 | .0074 | 1.0000 | | | | | | | |

zx( 2) .2408 -.2106 .0224 -.0032 .0003 -.0415 .4074 1.0000
zx( 3) -.2199 -.8117 .0083 -.0024 -.0008 -.0370 .0105 .4775 1.0000
zx( 4) -.4005 -.9142 .0044 -.0013 -.0020 -.0320 -.0871 .3466 .8910 1.0000
zy( 1) .0309 .0010 .2933 .2020 0.0000 .0146 .0824 .0653 .0230 .0112 1.0000
zy( 2) .0250 -.0030 .1991 -.2157 0.0000 -.0395 .0670 .0593 .0241 .0137 .3455 1.0000
zy( 3) .0097 -.0024 -.0525 -.7372 0.0000 -.0361 .0250 .0249 .0113 .0068 -.0021 .4628 1.0000
zy( 4) .0049 -.0010 -.2405 -.8965 -.0036 -.0301 .0101 .0121 .0059 .0031 -.1145 .3148 .8156 1.0000

ax bx ay by rs rn zx( 1) zx( 2) zx( 3) zx( 4) zy( 1) zy( 2) zy( 3) zy( 4)

For condition 1 : WE

Estimated Binormal ROC curve, with Lower and Upper Bounds of the Asymmetric Point-wise 95% Confidence Interval for True-Positive Fraction at a Variety of False-Positive Fractions:

| FPF | TPF | (Lower Bound, Upper Bound) |
|---|---|---|
| .005 | .4940 | ( .2083 , .7830 ) |
| .010 | .5565 | ( .2825 , .8051 ) |
| .020 | .6232 | ( .3731 , .8293 ) |
| .030 | .6638 | ( .4328 , .8449 ) |
| .040 | .6932 | ( .4776 , .8568 ) |
| .050 | .7162 | ( .5132 , .8665 ) |
| .060 | .7351 | ( .5427 , .8748 ) |
| .070 | .7512 | ( .5677 , .8822 ) |
| .080 | .7651 | ( .5893 , .8888 ) |
| .090 | .7774 | ( .6082 , .8947 ) |
| .100 | .7883 | ( .6249 , .9002 ) |
| .110 | .7982 | ( .6399 , .9053 ) |
| .120 | .8073 | ( .6535 , .9101 ) |
| .130 | .8155 | ( .6657 , .9145 ) |
| .140 | .8232 | ( .6769 , .9187 ) |
| .150 | .8303 | ( .6872 , .9226 ) |
| .200 | .8595 | ( .7283 , .9393 ) |
| .250 | .8817 | ( .7580 , .9523 ) |
| .300 | .8994 | ( .7809 , .9626 ) |
| .400 | .9263 | ( .8147 , .9774 ) |
| .500 | .9461 | ( .8398 , .9869 ) |
| .600 | .9614 | ( .8603 , .9929 ) |
| .700 | .9737 | ( .8786 , .9966 ) |
| .800 | .9838 | ( .8966 , .9987 ) |
| .900 | .9922 | ( .9168 , .9997 ) |
| .950 | .9959 | ( .9305 , .9999 ) |

Estimates of Expected Operating Points on fitted ROC curve, with lower and upper bounds of asymmetric 95% confidence interval along the curve for those points:

| Expected operating point ( FPF , TPF ) | Lower bound ( FPF , TPF ) | Upper bound ( FPF , TPF ) |
|---|---|---|
| ( .0035, .4646) | (0.0000, .0528) | ( .3949, .9251) |
| ( .1249, .8115) | ( .0086, .5427) | ( .5323, .9515) |
| ( .5142, .9485) | ( .3393, .9111) | ( .6864, .9722) |
| ( .5000, .9461) | ( .3575, .9159) | ( .6425, .9670) |

For condition 2 : WOE

Estimated Binormal ROC curve, with Lower and Upper

Bounds of the Asymmetric Point-wise 95% Confidence Interval for True-Positive Fraction at a Variety of False-Positive Fractions:

| FPF | TPF | (Lower Bound, Upper Bound) |
|---|---|---|
| .005 | .5036 | ( .2828 , .7234 ) |
| .010 | .5451 | ( .3352 , .7428 ) |
| .020 | .5898 | ( .3946 , .7646 ) |
| .030 | .6176 | ( .4325 , .7788 ) |
| .040 | .6381 | ( .4607 , .7897 ) |
| .050 | .6545 | ( .4832 , .7988 ) |
| .060 | .6683 | ( .5020 , .8066 ) |
| .070 | .6801 | ( .5181 , .8135 ) |
| .080 | .6906 | ( .5323 , .8197 ) |
| .090 | .7000 | ( .5449 , .8254 ) |
| .100 | .7086 | ( .5562 , .8307 ) |
| .110 | .7164 | ( .5665 , .8357 ) |
| .120 | .7236 | ( .5760 , .8403 ) |
| .130 | .7304 | ( .5847 , .8447 ) |
| .140 | .7367 | ( .5928 , .8489 ) |
| .150 | .7426 | ( .6003 , .8529 ) |
| .200 | .7682 | ( .6319 , .8705 ) |
| .250 | .7889 | ( .6565 , .8854 ) |
| .300 | .8066 | ( .6767 , .8983 ) |
| .400 | .8361 | ( .7089 , .9202 ) |
| .500 | .8608 | ( .7347 , .9383 ) |
| .600 | .8829 | ( .7572 , .9537 ) |
| .700 | .9036 | ( .7783 , .9670 ) |
| .800 | .9244 | ( .7999 , .9788 ) |
| .900 | .9472 | ( .8256 , .9893 ) |
| .950 | .9617 | ( .8440 , .9943 ) |

Estimates of Expected Operating Points on fitted ROC curve, with lower and upper bounds of asymmetric 95% confidence interval along the curve for those points:

| Expected operating point ( FPF , TPF ) | Lower bound ( FPF , TPF ) | Upper bound ( FPF , TPF ) |
|---|---|---|
| ( .0039, .4900) | (0.0000, .1498) | ( .4084, .8382) |
| ( .1264, .7280) | ( .0088, .5370) | ( .5350, .8688) |
| ( .6012, .8832) | ( .4233, .8421) | ( .7600, .9160) |
| ( .5000, .8608) | ( .3575, .8242) | ( .6425, .8918) |

'Plots' of the Fitted Binormal ROC Curves:

**********

| FPF | TPF for Condition 1 | TPF for Condition 2 |
|---|---|---|
| .005 | .4940 | .5036 |
| .010 | .5565 | .5451 |
| .020 | .6232 | .5898 |
| .030 | .6638 | .6176 |
| .040 | .6932 | .6381 |
| .050 | .7162 | .6545 |
| .060 | .7351 | .6683 |
| .070 | .7512 | .6801 |
| .080 | .7651 | .6906 |
| .090 | .7774 | .7000 |
| .100 | .7883 | .7086 |
| .110 | .7982 | .7164 |
| .120 | .8073 | .7236 |
| .130 | .8155 | .7304 |
| .140 | .8232 | .7367 |
| .150 | .8303 | .7426 |
| .200 | .8595 | .7682 |
| .250 | .8817 | .7889 |
| .300 | .8994 | .8066 |
| .400 | .9263 | .8361 |
| .500 | .9461 | .8608 |
| .600 | .9614 | .8829 |
| .700 | .9737 | .9036 |
| .800 | .9838 | .9244 |
| .900 | .9922 | .9472 |
| .950 | .9959 | .9617 |